Définir et éduquer l'esprit critique

Pasquinelli, E., Farina, M., Bedel, A., Casati, R.

Rapport produit dans le cadre des travaux du Work Package 1 / Projet EEC – Éducation à l'esprit critique (ANR-18-CE28-0018)

Juin 2020

Le présent rapport est le fruit de l'activité du Work package 1 (WP1) du projet EEC-Éducation à l'esprit critique (ANR-18-CE28-0018)

Le projet EEC – Éducation à l'esprit critique (ANR-18-CE28-0018)

L'esprit critique (EC) est au cœur des dernières politiques éducatives. Pourtant, si les projets éducatifs se multiplient, il existe encore très peu d'études scientifiques évaluant les méthodes proposées par les différents acteurs de l'éducation et de la formation. Ces méthodes ne s'appuient pas non plus sur les connaissances scientifiques existantes. Le présent projet vise à combler cette lacune en concevant et en testant des interventions pédagogiques pour le développement de l'esprit critique, et en établissant ainsi une approche de l'éducation de l'esprit critique fondée sur la preuve. De façon préalable, et grâce à la complémentarité des chercheurs et des enseignants de différentes disciplines, le projet vise à converger vers une définition plus précise de la pensée critique. Il vise en outre à produire une échelle d'évaluation de l'esprit critique. Enfin, il propose de diffuser ces ressources aux institutions et aux enseignants, ainsi qu'aux médias et au grand public.

Objectifs et travaux du *Work Package* 1 (WP1) – Analyse théorique. Définir et caractériser l'esprit critique.

Responsable : Roberto Casati (Institut Jean Nicod)

Auteurs du rapport
Elena Pasquinelli, Fondation *La main à la pâte*, Paris (France)
Mathieu Farina, Fondation *La main à la pâte*, Paris (France)
Audrey Bedel, Ecole Pratique des Hautes Etudes, Paris (France)
Roberto Casati, Institut Jean Nicod (CNRS/EHESS/ENS, DEC, Université PSL), Paris (France)

Ont en outre participé aux travaux du WP1 : Jie Gao (dans le cadre d'un stage du CogMaster PSL-P5-EHESS), Pierre Grialou (professeur des écoles), Nathanaël Jeune (étudiant en Master EdTech, CRI), Sarah Peyronnet (professeur des écoles), Calliste Scheibling-Sève (dans le cadre de sa thèse doctorale), Emmanuel Trouche (**Université Mohammed VI Polytechnique)** et Sacha Yesilaltay (doctorant, Institut Jean Nicod).

Ont été écoutés, lors des séances de préparation du rapport :

- Grégoire Borst, directeur du laboratoire Lapsydé, Université Paris-Descartes ;
- Hugo Mercier, CNRS, Institut Jean Nicod ;
- Florent Meyniel, Laboratoire Neurospin, CEA Saclay.

Le WP1 a présenté un rapport intermédiaire en juin 2019 au reste du consortium du projet et reçu ses commentaires et suggestions d'amélioration.
Le WP1 a eu pour objectif d'analyser la littérature existante, se focalisant sur la définition, les bases cognitives, l'éducabilité et la mesure de l'EC.

Caractérisation du concept d'« esprit critique » (EC) : des approches existantes à une définition minimaliste et opérationnelle

L'objectif principal du WP1 a consisté à fournir une caractérisation de l'EC qui permette de définir clairement (1) un besoin et (2) un objectif éducatif. Dans ce but, le groupe du WP1 a analysé la littérature concernant l'EC et ses différentes définitions. Au sein de cette littérature, le groupe a cherché à identifier les éléments à retenir pour une caractérisation opérationnelle de l'EC. Les approches actuelles s'étant révélées insatisfaisantes à plusieurs points de vue, nous nous sommes tournés vers d'autres approches, en dehors de la littérature spécifique sur l'EC, afin de remplir l'objectif principal.

Naturalisation de l'esprit critique

Le groupe WP1 a procédé à l'identification d'un nombre restreint de mécanismes et de fonctions cognitives impliqués dans la prise de décision, dans l'optimisation des décisions et, notamment, dans l'évaluation (critique) de l'information nécessaire à la prise de décision. Ces fonctions constituent des potentiels « *building blocks* » cognitifs de l'esprit critique et font donc partie d'une approche cognitive de l'EC. Ces *building blocks* (1) servent de base pour évaluer l'information (de première et de seconde main) et (2) permettent d'ajuster la confiance dans l'information ainsi évaluée, afin de prendre une décision.
Le groupe a fourni un effort particulier pour identifier des fonctions minimalistes et pour tracer un cadre écologiquement valide pour l'EC.
Cependant, cet objectif reste au niveau d'ébauche : la liste des mécanismes et des fonctions identifiés n'est pas exhaustive, leurs relations n'étant pas encore complètement spécifiées ou réunies dans un modèle.

Éducabilité à l'EC

Le groupe du WP1 s'est concentré sur la problématique du transfert des capacités et sur l'opportunité d'insérer l'enseignement de la pensée critique dans un contexte disciplinaire. Dès que l'élève fait preuve de pensée critique dans le domaine des mathématiques, est-il alors capable de transférer cette habileté à l'apprentissage de l'histoire ou à l'évaluation des informations ? Cette problématique est fondamentale pour l'éducation de l'EC car le but final d'une telle éducation ne peut être que celui d'une amélioration de la prise de décision dans la vie de tous les jours, qui présente des contextes très différents. La possibilité d'acquérir des capacités générales indépendantes du contexte et du contenu est très débattue. Nous avons parcouru la littérature concernant cette problématique pour nous concentrer ensuite sur celle plus spécifique sur les manières de transférer des capacités acquises d'un contexte à un autre, plus ou moins éloigné.
Nous avons finalement répertorié plusieurs synthèses de la littérature concernant l'impact des méthodes actuelles pour l'éducation de l'EC et la transférabilité des résultats obtenus. Il est apparu que la littérature concernée était extrêmement hétérogène, ce que rendait sa synthèse difficile sans prendre le risque de comparer des notions incomparables. Parmi les méthodes censées être dédiées à l'éducation de l'EC, certaines sont en effet plutôt tournées vers les capacités de compréhension des textes, vers la méthode scientifique, d'autres vers la reconnaissance des biais, l'acquisition des outils mathématiques et scientifiques permettant de contrecarrer certains biais, etc.

Résultats de l'étude

L'étude a résulté sur une définition minimaliste, opérationnelle et cognitivement réaliste de l'EC, et des principes permettant de guider la pratique pédagogique et l'évaluation. En résumé :

- l'EC est défini comme l'ensemble des capacités qui permettent d'évaluer la qualité épistémique des informations disponibles en vue d'une prise de décision, et de calibrer correctement sa confiance en ces informations selon les résultats de l'évaluation ;
- l'EC n'est pas uniquement un idéal à atteindre, mais il fait partie du bagage cognitif naturel de tout individu, présent dès l'enfance. Nous faisons preuve d'EC quand nous sélectionnons nos sources sur la base de critères ou d'indices de fiabilité, quand nous estimons la plausibilité et la pertinence d'une information, quand nous ressentons ou exprimons notre confiance dans les informations à notre disposition et dans nos choix ou décisions. Les enfants possèdent ces capacités et les utilisent de manière souvent inconsciente. Nous considérons ces capacités comme constitutives de l'EC naturel ;
- les capacités naturelles de l'EC présentent des limites structurelles : elles ne sont pas sans faille, et les critères ou indices utilisés ne sont pas nécessairement adaptés à des contextes et contenus qui demandent un EC plus sophistiqué ;
- l'éducation de l'EC a pour but d'outiller les capacités de l'EC naturel avec des critères de plus en plus sophistiqués. Son premier objectif consiste à mettre tout citoyen en condition pour évaluer correctement les sources et les contenus d'information d'usage courant. Nous parlons pour cela d'EC « avancé ». Nous distinguons l'EC « avancé » de l'EC « expert », qui s'exerce sur des contenus moins courants et qui utilise des critères encore plus sophistiqués, des connaissances spécialisées voire professionnelles ;
- afin d'atteindre ses objectifs, l'éducation de l'EC se soucie de mettre en place les meilleures stratégies pour favoriser le transfert des capacités et outils acquis dans de nombreux domaines de connaissance et contextes ;
- l'évaluation de l'EC vise à mesurer la capacité à mettre en place les outils les plus adaptés pour mesurer la qualité épistémique de l'information et également la capacité à calibrer sa confiance en conséquence.

DEFINIR ET EDUQUER L'ESPRIT CRITIQUE 1

Pasquinelli, E., Farina, M., Bedel, A., Casati, R. 1
OBJECTIFS ET TRAVAUX DU *WORK PACKAGE* 1 (WP1) – ANALYSE THEORIQUE. DEFINIR ET CARACTERISER L'ESPRIT CRITIQUE. 2

1. INTRODUCTION 7

1.1 POURQUOI L'EC EST-IL IMPORTANT ? 7
1.2 MAIS COMMENT ? 8
1.3 METHODE 9

2. VERS UNE DEFINITION OPERATIONNELLE DE L'EC 10

2.1 ANALYSE DES APPROCHES EXISTANTES. CONSTAT DE L'ABSENCE D'UNE DEFINITION CONSENSUELLE 10
2.1.1 Un peu d'histoire. Approches philosophiques 11
2.1.2 Les approches psychologiques 24
2.2 DES RAISONS D'INSATISFACTION PAR RAPPORT AUX APPROCHES EXISTANTES DE L'EC 40
2.2.1 Des définitions trop vastes pour être opérationnelles 40
2.2.2 Des définitions trop vagues, qui peuvent donner lieu à des malentendus 41
2.2.3 Des définitions qui font référence à des concepts ambigus 42
2.2.4 Des définitions trop peu réalistes 44
2.3 L'EC, QU'EST-CE QUE C'EST ? 45

3. LES BASES COGNITIVES DE L'EC 47

3.1 NATURALISER L'EC 47
3.2 LES BASES COGNITIVES DE L'EC : LA VIGILANCE EPISTEMIQUE 49
3.2.1 Les mécanismes de la vigilance épistémique 49
3.2.1.1 Les mécanismes de la confiance sélective chez l'enfant (vigilance dirigée sur les sources) 50
3.2.1.2 L'épistémologie naïve de l'enfant (vigilance dirigée sur les contenus) 51
3.2.2 Les mécanismes de la vigilance épistémique ne sont pas infaillibles 53
3.2.3 Des formes sociales et culturelles de vigilance épistémique permettent de remédier aux limites de la vigilance épistémique naturelle. Naissance de la vigilance épistémique étendue 53
3.2.4 Expliquer la crédulité, les opinions irrationnelles et les effets de persuasion 54
3.2.4.1 Nous ne sommes pas aveuglément crédules, mais plutôt conservateurs 54
3.2.4.2 De bonnes raisons pour accepter les idées d'autrui 55
3.2.4.3 L'influence des contenus sur l'acceptation des idées 55
3.2.4.4 Le rôle de l'appartenance au groupe 56
3.2.4.5 Le rôle du prestige de la source 57
3.2.5 L'incertitude comme moteur pour changer de position et suivre l'avis d'autrui 58
3.2.6 La vigilance épistémique. Conséquences pour l'éducation 59
3.3 LES BASES COGNITIVES DE L'EC : LA SENSIBILITE METACOGNITIVE, LA CONFIANCE 61
3.3.1 Aspects de base de la métacognition 62
3.3.2 La métacognition implicite 63
3.3.3 Trajectoires développementales de la métacognition 63
3.3.3.1 Les mécanismes précoces de la sensibilité métacognitive 65
3.3.4 Confiance et cerveau 67
3.3.4.1 Confiance et cerveau bayésien 67
3.3.4.2 Confiance, sensibilité à l'erreur et cerveau prédictif 69
3.3.5 Forces et limites de la confiance métacognitive : une question de calibration 70
3.3.6 Biais métacognitifs entre sousconfiance et surconfiance 70
3.3.6.1 La surconfiance dans les jugements et l'effet dit « Dunning-Kruger » 71
3.3.6.1.1 L'effet Dunning-Kruger dans des tâches de raisonnement 72
3.3.6.2 L'illusion de profondeur explicative 73
3.3.7 Pourquoi la sensibilité métacognitive pourrait-elle être mal calibrée ? 74
3.3.7.1 Les heuristiques à l'origine des illusions métacognitives : le problème de la nature indirecte des indices 74
3.3.8 Stratégies qui permettent d'améliorer la qualité des jugements métacognitifs 75
3.3.9 La confiance métacognitive. Conséquences pour l'éducation de l'EC 77
3.4 CHERCHER A COMPRENDRE LA NATURE DE NOS ERREURS 79
3.4.1 Le programme « heuristiques et biais » (H & B) et la nature des erreurs 79
3.4.2 Critiques du programme « H & B » 80
3.4.3 L'approche « évo-éco » : remettre nos intuitions et instincts dans une lumière positive 80
3.4.3.1 Un exemple d'une approche « évo-éco » à l'explication de décisions et comportements 81
3.4.3.2 Un exemple d'approche « évo-éco » pour l'explication des illusions perceptives 82
3.4.3.3 À l'origine de nos erreurs 83
3.4.4 Normes externes et normes internes : retour à l'EC 84
3.4.4.1 « Ne pas jeter le bébé avec l'eau du bain » : intérêt d'une approche H & B dans le cadre d'une approche cognitive de l'EC 85
3.4.4.2 Que faire avec nos biais et nos erreurs dans le cadre d'une approche éducative de l'EC ? 85
3.4.4.3 Inhiber ? 87
3.4.4.4 Prendre conscience des biais les plus communs et des circonstances dans lesquelles ils peuvent nous mettre en difficulté 89

3.4.5 Heuristiques, erreurs récurrentes et biais. Des indications pour l'éducation de l'EC *91*
3.5 SYNTHESE ET REMARQUES GENERALES 92

4. EDUCABILITE DE L'ESPRIT CRITIQUE 95

4.1 ÉDUQUER DES CAPACITES GENERALES 95
4.2. L'EC EST-IL UNE CAPACITE GENERALE ? 95
4.2.1 Le problème de l'intelligence ne se pose pas *96*
4.2.2 Heuristiques générales : oui ou non ? *97*
4.2.3 Atouts des experts *97*
4.2.4 Le cerveau n'est pas un muscle *99*
4.2.5 Savoir et savoir appliquer *99*
4.2.6 Stratégies (apprendre à apprendre) *100*
4.3. APPRENDRE A PENSER DE FAÇON (PLUS) CRITIQUE : UN FAUX ESPOIR ? 101
4.3.1 Le pessimisme de Daniel Willingham *101*
4.3.2 L'optimisme de Diane Halpern *103*
4.3.2.1 Une méta-analyse 104
4.3.2.2 Une large intervention 105
4.3.2.3 Une étude d'impact d'une méthode spécifique 105
4.3.2.4 Des études avec mesure de transfert à distance 106
4.3.3 Une nouvelle synthèse : Abrami et al. (2015) *106*
4.4. DIFFICILE NE VEUT PAS DIRE IMPOSSIBLE : REALITE ET DIFFICULTES DU TRANSFERT 108
4.4.1 La notion de transfert est fondamentale pour l'apprentissage *109*
4.4.1.1 Transfert et analogie 110
4.4.1.2 Similitudes et analogies : mais à quel niveau ? 111
4.4.1.3 Limites de l'analogie 113
4.4.2 Avantages de la variation des contextes *113*
4.4.3 Avantages de l'explicitation *114*
4.4.4 Voie basse et voie haute *115*
4.4.5 Beaucoup de pratique, mais pas n'importe laquelle *116*
4.4.6 La métacognition et les apprentissages métacognitifs *117*
4.4.7 L'argumentation comme support de l'esprit critique social et comme outil pédagogique *118*
4.4.7.1 L'approche argumentative (1) 118
4.4.7.2 L'approche argumentative (2) 119
4.4.7.2.1 Argumentation et évaluation des arguments d'autrui 120
4.5. SYNTHESE ET CONSIDERATIONS PRATIQUES POUR LES PEDAGOGIES DE L'EC 121
4.5.1 Conseil pratiques pour favoriser le transfert de l'EC *122*

5. AVIS PEDAGOGIQUES : CHOIX A FAIRE ET A NE PAS FAIRE DANS L'EDUCATION DE L'EC 124

5.1 À NE PAS FAIRE 124
5.1.1 Ne pas réduire l'éducation de l'EC à une éducation aux médias de l'information *124*
5.1.2. Faire attention à ne pas donner l'impression que ce qui compte dans un débat argumentatif est de bien présenter ses arguments *125*
5.1.3. Ne pas se limiter à des thèmes chauds *126*
5.1.4. Ne pas limiter l'enseignement de l'EC à une prise de connaissance des biais cognitifs, et éviter les listes de biais *127*
5.1.5. Ne pas se limiter à des enseignements implicites, non assumés de l'EC, ni se limiter à un enseignement abstrait et théorique de celui-ci *127*
5.2 À FAIRE 129
5.2.1 Outiller l'EC naturel. Comment choisir les critères à enseigner ? *130*
5.2.2 Où trouver ces critères dans les programmes de l'éducation française ? *131*
5.2.3 Exercer la confiance métacognitive de façon explicite. Se placer sur un continuum de confiance et réfléchir à sa signification *132*
5.2.4 Éléments additionnels pour parvenir à une juste évaluation de la qualité de l'information et un calibrage de la confiance *133*
5.2.4.1 Acquérir une plus vaste et profonde base de connaissances 133
5.2.4.2 Prendre conscience des situations où l'exercice de l'EC avancé est requis et des circonstances plus à même de nous mettre en difficulté 134
5.2.5 Tester *135*

1. Introduction

1.1 Pourquoi l'EC est-il important ?

L'esprit critique (EC) est la saveur du jour, dans l'éducation et dans le débat public. L'éducation de l'EC est invoquée par les enseignants, les parents, les ministères de l'éducation et les organisations internationales. Ses vertus supposées vont de la promotion de la capacité des étudiants – en tant que futurs citoyens et travailleurs – à l'application de « leurs connaissances dans des circonstances nouvelles et évolutives » (Howels 2018) et à celle de permettre à des informations pertinentes et crédibles d'être « sélectionnées, interprétées, digérées, évaluées, apprises et appliquées » (Halpern 2013). L'EC est aujourd'hui considéré comme plus important que jamais pour relever les défis personnels et sociétaux, pour la résolution de problèmes et pour devenir des décideurs plus compétents.

« *Bien que la capacité de penser de manière critique ait toujours été importante, c'est une nécessité vitale pour les citoyens du XXIe siècle* » (Halpern 2013).

Parmi les circonstances qui rendent l'EC si pertinent à l'heure actuelle, les tendances de l'utilisation des médias et des réseaux sociaux jouent un rôle important : à l'ère de l'info-obésité, beaucoup évoquent les dangers des fausses nouvelles et des attitudes post-vérité, associés à un manque de compétences de base pour juger correctement les informations (Whitworth 2009 ; voir Acerbi, 2019, pour une critique des vues simplistes de l'impact des fausses nouvelles) tandis que d'autres insistent sur la difficulté de se protéger de nos propres biais (Pronin & Ross 2002). Simultanément attaqués de l'extérieur et faibles à l'intérieur, nous risquons de prendre de mauvaises décisions, de soutenir des opinions simplistes qui ne seraient pas étayées par des preuves, et de devenir crédules ou, inversement, uniformément sceptiques.

De nombreux chercheurs travaillant dans le domaine de l'esprit critique déplorent le mauvais état de ce dernier chez la plupart des adultes et des enfants instruits. Par exemple, Halpern (1998) signale que de nombreux adultes – sinon la plupart – ne pensent pas de manière critique dans de nombreuses situations. Elle observe qu'un grand nombre de personnes ont des croyances irrationnelles, par exemple quant aux phénomènes paranormaux : elles font des choix irrationnels et ne cherchent pas de preuves pertinentes. Kuhn (1999) décrit la voie de développement des compétences métacognitives et épistémologiques qu'elle juge cruciales pour l'EC – à savoir des compétences liées à l'appréciation du rôle de la preuve dans l'acceptation ou la mise à jour des théories et des croyances – pour souligner que ces capacités sont presque aussi déficientes chez l'adulte que chez l'enfant.

Beaucoup d'autres concluent également que la plupart des adultes ne disposent pas de compétences de raisonnement de base et qu'ils sont la proie d'inférences qui « v*iolent fréquemment les principes de la statistique, de l'économie, de la logique et de la méthodologie scientifique de base* » (Nisbett 2015).

Les traditions psychologiques, telles que le programme dit des « heuristiques et biais », fondent leur pessimisme sur l'observation selon laquelle les individus affichent une multitude de biais violant la structure de la rationalité. « *Par exemple, les gens affichent un biais de confirmation, testent les hypothèses de manière inefficace, affichent des incohérences de*

préférence, ne calibrent pas correctement les degrés de croyance, projettent trop leurs propres opinions sur les autres, combinent les probabilités de manière incohérente et permettent aux connaissances antérieures d'être impliquées dans le raisonnement déductif » (Nisbett 2015 ; pour les résumés de la littérature, voir Evans 2012 ; Gilovich, Griffin & Kahneman 2002 ; Kahneman 2002 ; Shafir & LeBoeuf 2002 ; Stanovich 1999, 2009). La liste de biais cognitifs est longue : négligence du taux de base, effets de cadrage, biais de représentativité, biais d'ancrage, biais de disponibilité, biais de résultat et effets de vivacité, pour n'en nommer que quelques-uns (Stanovich & Stanovich 2010). Selon une certaine vision de la rationalité humaine, les êtres humains sont davantage des chercheurs de modèles et des conteurs qu'ils ne sont naturellement des penseurs critiques. « *En effet, comme le ballet, la pensée critique est une activité très recherchée. La course est naturelle ; la danse de boîte de nuit l'est moins ; mais le ballet est quelque chose que les gens ne peuvent bien faire qu'avec de nombreuses années de formation douloureuse, coûteuse et dédiée. L'évolution ne voulait pas que nous marchions jusqu'au bout de nos orteils, et quoi que puisse dire Aristote, nous n'étions pas conçus pour être tout aussi critiques. L'évolution ne gaspille pas d'efforts pour rendre les choses meilleures qu'elles ne devraient l'être…* » (van Gelder, 2005).

Les partisans de l'EC affirment que l'éducation à l'EC devrait viser à remédier à ce sombre état de choses, notamment en faisant passer les dispositions de pensée de l'irrationnel au rationnel. « *En bref, une grande partie de la justification des interventions éducatives pour changer les dispositions de la pensée découle d'une supposition tacite selon laquelle les dispositions de la pensée critique font de l'individu une personne plus rationnelle – ou, comme le soutient Sternberg (2001, 2005) – une personne plus sage et moins stupide. Ainsi, la justification normative de la stimulation de la pensée critique est qu'elle est le fondement de la pensée rationnelle* » (Stanovich & Stanovich 2010).

1.2 Mais comment ?

Comment « passer de l'irrationnel au rationnel » ? Aucune réponse simple n'existe encore pour cette question.

Premièrement, parce que le terme « esprit critique » a plusieurs significations différentes. Cette polysémie donne lieu à une multiplicité d'interventions pédagogiques visant à développer l'EC, mais les contenus des interventions se recoupent rarement. Ce qui devrait être inclus dans un programme éducatif visant à renforcer la pensée critique n'est, par conséquent, pas encore clair. Les méthodes existantes sont à la fois hétérogènes et souvent très larges, englobant une grande variété de compétences et de critères. Les outils d'évaluation de l'EC souffrent du même problème. De plus, il existe peu d'outils qui se veulent adaptés spécifiquement aux enfants.

Deuxièmement, et c'est l'élément clé de notre diagnostic, parce que la plupart des approches actuelles de l'EC ne traitent pas la question des éléments constitutifs cognitifs naturels qui rendent l'EC possible en premier lieu.

Un effort est donc nécessaire pour bâtir une théorie de l'EC pouvant donner lieu à des résultats opérationnels (c'est-à-dire des méthodes pédagogiques et des outils d'évaluation associés). Une telle théorie devrait être basée sur une définition restreinte (au sens de spécifique) de l'EC. De cette façon, il serait possible de développer des actions ciblées ainsi que des outils d'évaluation dédiés, et de comparer les résultats des différentes actions pédagogiques.

Nous postulons également qu'une théorie de l'EC devrait être informée par les connaissances actuelles en science cognitive, c'est-à-dire par une compréhension des fonctions cognitives impliquées dans l'EC, de leur trajectoire de développement et de leurs limites. L'avantage de cette approche consiste à fournir un terrain objectif sur lequel des visées pédagogiques puissent être définies en termes de développement ou d'outillage de fonctions naturelles préexistantes.

1.3 Méthode

Nous analysons d'abord les approches rencontrées dans la littérature sur l'EC. Ce travail nous permet de construire une nouvelle définition de l'EC, dans le sillage de la tradition des études sur l'EC existantes, mais compatible avec l'objectif opérationnel.

Ensuite, en nous basant sur la définition proposée et à travers une analyse de la littérature en psychologie du développement, en neurosciences cognitives et en psychologie évolutive, nous identifions un nombre limité de « *building blocks* » cognitifs qui rendent l'EC possible. Nous explorons leur trajectoire de développement et mettons en évidence leurs limites.

Une meilleure compréhension des fondements naturels est une condition nécessaire pour proposer des stratégies pédagogiques scientifiquement fondées pour le développement de l'EC. Nous proposons donc un ensemble de principes pédagogiques compatibles avec la caractérisation des capacités naturelles de l'EC, tout en présentant aussi leurs limites et leur marge de progression.

Une attention particulière est accordée à la question du transfert des compétences liées à l'EC. Il s'agit d'une préoccupation majeure pour les éducateurs, d'autant plus que la possibilité d'éduquer des capacités dites générales comme l'EC a été à plusieurs reprises remise en question.

2. Vers une définition opérationnelle de l'EC

2.1 Analyse des approches existantes. Constat de l'absence d'une définition consensuelle

Dans la littérature académique, l'EC est avant tout un **objectif éducatif**, c'est-à-dire un état désirable qui justifie un effort éducatif particulier (Hitchcock 2018, Lai 2012[1]). Toutefois, le terme « esprit critique » a reçu et n'a de cesse de recevoir une variété de définitions faisant référence à différentes traditions, notamment en philosophie et en psychologie de l'éducation, chacune comportant des vues différentes. **Une certaine confusion persiste donc quant à la nature de l'esprit critique, en dépit du fait que ce terme dispose d'une côte de plus en plus importante.**
Selon certains auteurs, en dépit des différences apparentes, les définitions existantes convergeraient autour d'un même concept ou « cluster » conceptuel : celui d'une **attitude réflexive**, permettant de décider quelles croyances accepter ou réfuter ; d'une **forme d'analyse active et attentive des croyances et des preuves en leur faveur** ; d'un **processus discipliné** qui permettrait d'analyser, de synthétiser, de conceptualiser, d'évaluer, d'utiliser l'information ; d'une **mobilisation de bonnes raisons**, de critères ; de la **capacité de s'autocorriger** ; d'une forme de **pensée ouverte, mais aussi autonome** ; d'une **attitude inquisitive** et de **capacités argumentatives** (voir, par exemple, Ennis 2016[ii]). D'autres auteurs (voir, par exemple, Halpern 2013[3]) soutiennent qu'au cœur de toutes les définitions se trouvent des **fonctions** telles que raison, logique, jugement, métacognition, réflexion, questionnement, et d'autres processus mentaux ayant trait au raisonnement qui permet d'arriver à une solution ou conclusion de manière justifiée.
D'autres encore reconnaissent que le consensus concerne l'**objectif d'éduquer** l'esprit critique plutôt que la définition elle-même du terme. L'absence de consensus deviendrait particulièrement évidente au moment de passer de la théorie – et de définitions très générales et inclusives – à la pratique, qui nécessite de décider sur quoi concentrer les actions éducatives (Bailin et al. 1999[4]). Exemplaire, à ce niveau, est la définition « consensuelle » produite, dans le cadre de l'effort de recherche de consensus, par un groupe de philosophes, chercheurs en sciences sociales, éducateurs et scientifiques (physiciens), sous l'impulsion de l'*American Philosophical Association*, à la fin des années 1980. Pour cela, les participants ont suivi différentes étapes structurées de recherche de consensus, selon le modèle connu comme « processus Delphi » : à chaque étape, des définitions sont fournies, commentées, et un terrain commun est recherché jusqu'à aboutir à une définition finale et majoritairement consensuelle des capacités et des dispositions propres à la pensée critique.

La définition qui a émergé s'applique cependant uniquement à un « penseur idéal » :

"*We understand critical thinking to be the purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based. CT is essential as a tool of inquiry. As such, CT is a liberating force in education and a powerful resource in one's personal and civic life. While not synonymous with good thinking, CT is a pervasive and self-rectifying human phenomenon. The ideal critical thinker is habitually inquisitive, well-informed, trustful of reason, open-minded, flexible, fair-minded in evaluation, honest in facing personal biases, prudent in making judgments, willing to reconsider, clear about issues, orderly in complex matters, diligent in seeking relevant information, reasonable in the selection of criteria, focused in inquiry, and*

persistent in seeking results which are as precise as the subject and the circumstances of inquiry permit. Thus, educating good critical thinkers means working toward this ideal. It combines developing CT skills with nurturing those dispositions which consistently yield useful insights and which are the basis of a rational and democratic society." (Facione 1990[5])

Les auteurs considèrent que l'effort d'éduquer pour améliorer la pensée critique reste un objectif valable, et que l'« idéal » ainsi défini devrait servir de guide afin d'imaginer une éducation de l'EC. Cependant, la nature idéalisée de la définition rend difficile d'établir ce qu'il est possible d'atteindre et ce qui ne l'est pas.
De plus, la définition est assez large, au sens où elle fait référence à une grande variété de compétences et de dispositions, ce qui rend encore plus difficile d'imaginer le passage à l'opérationnalité. Ces capacités et dispositions ne sont pas non plus nécessairement spécifiques à la pensée critique : elles peuvent par exemple être impliquées dans la communication efficace, le raisonnement logique, la résolution de problèmes. Les capacités listées comme nécessaires à l'EC incluent en effet des capacités liées à la compréhension d'un message, des capacités de communication, d'expression et de justification, ainsi que d'autres encore, comme l'analyse et l'évaluation d'arguments, l'inférence, l'explication, l'autorégulation. De manière synthétique, l'EC comprend six dimensions ou catégories de capacités : capacités d'interprétation, d'analyse, d'évaluation, d'inférence, d'explication et d'autorégulation. En plus de ces capacités, le penseur critique idéal possèderait des dispositions à utiliser ces capacités et donc un « esprit », des inclinaisons, des vertus. Mais, justement, il ne s'agit que d'un idéal : le panel n'a pas pu atteindre un consensus sur l'opportunité d'inclure ces dispositions dans la définition de l'EC.
Conscient de cette situation, le psychologue de l'éducation Daniel Willingham – particulièrement intéressé par les aspects pragmatiques de l'éducation de l'EC – a proposé une définition de sens commun de l'EC. Selon celle-ci, l'EC est une forme de pensée efficace, c'est-à-dire qui amène à des bonnes solutions à un problème, qui s'exerce de manière volontaire, non automatique, et qui ne repose pas sur des solutions prêtes à l'emploi déjà stockées en mémoire (Willingham 2007, 2019[6]). Cependant, cette définition est, même selon Willingham, extrêmement générale – elle équivaut à « bien penser » – et doit en réalité être déclinée de manière spécifique dans chaque contexte et pour chaque nouveau contenu.

2.1.1 Un peu d'histoire. Approches philosophiques

Retracer, même sommairement, l'histoire des différentes approches de l'EC nous permettra de mieux apprécier les similitudes et différences entre les diverses définitions, entre les différents outils de mesure de l'EC existants et entre les formes d'intervention proposées pour favoriser le développement de l'EC.

Certains font remonter le concept d'EC à l'Antiquité, à Aristote et aux Sceptiques notamment. Dans cette reconstruction, l'EC serait au premier chef une forme de pensée respectant les règles de la logique. La « formation de l'esprit » serait aussi au centre des *Règles pour la direction de l'esprit* de Descartes et de la théorie du raisonnement de Hobbes ; on la retrouverait en fondement de la science moderne chez Francis Bacon et sa critique des idoles (*The Advancement of Learning)*, chez Robert Boyle et son modèle de pensée sceptique (*Sceptical Chymist*) (Paul, Elder, Bartell 1997 ; Foundation for critical thinking http://www.criticalthinking.org/pages/a-brief-history-of-the-idea-of-critical-thinking/4[7]). Elle serait plus généralement une composante de la pensée scientifique s'étant affirmée aux XVIIe et XVIIIe siècles, le siècle des Lumières.

Cependant, ce n'est que plus récemment, à partir des années 1970-1980, que la formation de l'EC est devenue un thème de réflexion générale en éducation, notamment grâce à un groupe de philosophes états-uniens qui ont repris l'idée d'un lien fort entre pensée critique et éducation (Lai 2011 ; Lewis & Smith 1993). Cette tradition est fortement influencée par le

travail séminal de John Dewey, philosophe de l'éducation très actif au cours des années 1930. L'EC devient, en effet, dans les années 1930, un objectif pédagogique des écoles qui adhèrent au programme deweyen (*Progressive Education Association*).

Le rôle fondateur de John Dewey (1859-1952)

En tant qu'éducateur, John Dewey s'intéresse fortement à la psychologie et à l'étude du développement de la pensée (Boisvert 2015 ; Dewey 1933, 1997). L'école est pour lui un lieu idéal pour étudier ce développement dans un environnement contrôlé, et pour permettre aux capacités naturelles de l'enfant de se développer dans des directions plus variées et complexes que celles sollicitées par l'expérience « naturelle », non guidée. Dans le modèle éducatif deweyen, l'objet de l'éducation intellectuelle (par opposition à l'éducation morale) est la pensée réflexive, qui correspond à la pensée critique : la pensée réflexive coïncide en effet avec l'activité consistant à évaluer le degré de probabilité qu'une croyance soit vraie, afin de l'accepter ou de la refuser, sur la base d'une analyse de ses fondements, donc des faits qui la soutiennent.

"Active, persistent and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it, and the further conclusions to which it tends, constitutes reflective thought." (Dewey 1933[8])

La pensée réflexive implique deux mouvements : d'une part, un acte d'investigation et de recherche : l'étude des faits et l'examen et la révision des preuves, la mise au point de différentes hypothèses et de leurs implications, la comparaison des hypothèses entre elles et avec les résultats de l'observation et des faits ; d'autre part, l'adoption d'une attitude sceptique : pour chercher, il faut d'abord douter, hésiter, être perplexe. Une telle attitude s'oppose à celle qui consiste à accepter une croyance sur la base de la tradition, de l'instruction, de l'imitation – donc de tout ce qui a trait à l'autorité – mais aussi sur la base de l'avantage qu'on espère en tirer, ou de réactions purement émotionnelles[9].

La définition de Dewey ne se focalise pas sur l'aboutissement de l'éducation à l'EC, sur le comportement désirable, mais sur les processus impliqués – suspendre judicieusement le jugement, déterminer la nature du problème à résoudre, évaluer les arguments avant de prendre une décision ou exprimer une opinion (au besoin) – et elle est fortement liée au raisonnement scientifique. (Afin d'illustrer sa définition, Dewey utilise plusieurs exemples et contre-exemples[10].) Selon lui, l'être humain n'est pas naturellement critique et le fait que nous soyons des organismes intelligents ne nous met pas à l'abri de la propagation de l'erreur et de l'accumulation de conceptions erronées. La crédulité dont l'être humain est affecté lui est dictée par la tendance à accepter comme vraie la première idée qui lui vient à l'esprit. Le fait de douter (le premier pas vers la pensée critique selon Dewey) a en effet des coûts : il est pénible et nous met dans une condition inconfortable d'incertitude. De surcroît, non seulement la nature humaine, mais la société elle-même, n'est pas encore suffisamment rationnelle, et elle instille donc de mauvaises habitudes de pensée par le biais de l'autorité, de l'instruction et de l'imitation. Cependant, l'enfant humain est tout aussi naturellement curieux et forme spontanément des idées, des suggestions ; il est, de plus, persévérant. L'éducation de la pensée critique se fonde donc sur l'exercice de ces capacités naturelles et sur leur transformation – à travers l'entraînement de l'esprit – en habitudes de pensée. Cette éducation constitue tout l'objectif de l'éducation intellectuelle. Elle doit aboutir à produire un esprit

discipliné, qui se sert de la pensée réflexive de manière experte, c'est-à-dire, plus spécifiquement, un esprit qui sache :

(a) identifier et définir une difficulté ;
(b) développer une suggestion pour sa solution et raisonner sur les conséquences de la suggestion (grâce aux capacités de déduction et d'induction) ;
(c) conduire d'autres observations et expériences pour accepter ou rejeter la suggestion (les étapes formelles de la pensée).

Il s'agit aussi et surtout de développer un esprit qui ait une disposition à utiliser ces capacités, et ceci dès le plus jeune âge. La science, en particulier la méthode expérimentale, utilise toutes les opérations de la pensée réflexive, et elle le fait de manière systématique, élaborée et spécialisée. La méthode scientifique expérimentale est alors le modèle à adopter en éducation, à appliquer à n'importe quel problème, matière, croyance, dès l'école primaire. Toutefois, cela ne veut pas dire que l'enseignement scientifique est particulièrement formateur ou doit avoir une place à part dans l'instruction. Au contraire, la dimension de la méthode scientifique doit entrer dans tous les aspects de l'éducation : même les disciplines manuelles, graphiques, occupationnelles, les activités les plus diverses, devraient et pourraient être utilisées pour présenter aux élèves des problèmes à résoudre à travers l'expérimentation, et de manière à susciter la nécessité d'acquérir des corpus de connaissances. Ces connaissances scientifiques plus spécialisées font ensuite l'objet d'une éducation spécifique, dans un deuxième moment. La pensée expérimentale, réflexive et critique devient ainsi le but de l'éducation, sans se constituer en enseignement à part. Dans cette acception, la formation de l'esprit critique sert d'ailleurs autant à la science, en tant qu'entreprise de connaissance par excellence, qu'au bien vivre commun, car adopter une croyance plutôt qu'une autre affecte à la fois nos comportements et d'autres croyances.

De Edward Glaser à Robert Ennis

L'approche de l'EC développée aux États-Unis dans les années 1980 est également fortement influencée par le travail d'Edward Glaser, un éducateur américain. En 1941, Glaser publie un livre puis un test de l'EC qui est encore utilisé aujourd'hui (Glaser 1941[11], Watson 1980). Pour lui, les trois piliers de l'esprit critique sont :

- l'attitude réflexive ;
- la connaissance des méthodes du raisonnement et de l'investigation logiques ou rationnelles ;
- la capacité à mettre en place ces méthodes.

Attitudes et compétences sont donc également considérées comme nécessaires dans cette tradition afin de favoriser l'EC.

Le terrain est enfin préparé par des réflexions dans le domaine éducatif, concernant les capacités qui devraient faire l'objet d'éducation. Nous pouvons rappeler la taxonomie des capacités cognitives proposée par Bloom en 1965. Celle-ci place l'acquisition de connaissances au niveau le plus bas de sa pyramide de capacités à faire acquérir. Viennent ensuite la compréhension, qui demande d'aller au-delà de la connaissance, l'application de ce qui est compris, l'analyse ou l'évaluation critique de ce qui est compris et appliqué, la synthèse des connaissances de domaines différents et, enfin, l'évaluation critique des connaissances qui ont été analysées et synthétisées. Le terme « *critical thinking* » n'est pas utilisé, mais les capacités cognitives indiquées font partie du groupe des capacités « d'ordre supérieur », qui incluent entre autres pensée critique, résolution de problèmes et métacognition.

Ainsi, en 1962, le philosophe Robert Ennis propose une première définition du concept d'EC comme « *the correct assessment of statements* » (Ennis 1962). Ensuite, il développe cette définition jusqu'à inclure un grand nombre de capacités et de dispositions qui définissent le penseur critique (Lewis & Smith 1993[12] ; voir, par exemple, Ennis 1962, 1964, 1989, 1991, 2011, 2016).
Les capacités de l'EC se divisent en nécessaires et accessoires, mais aussi en générales et spécifiques (à un certain contenu). Les capacités nécessaires comprennent :

- la **clarification**, niveau basique : il s'agit notamment de se focaliser sur la question à laquelle nous voulons répondre, d'analyser les arguments, de poser ou de se poser les questions nécessaires pour clarifier le problème ;
- l'**évaluation des informations**, nécessaire pour prendre une décision : juger la crédibilité d'une source, sur la base de critères comme la réputation, la présence de justification, la cohérence, etc. ; mener des observations ou juger la qualité d'une observation, toujours sur la base de ces critères ;
- l'**inférence** : déduction logique, généralisation, émettre des jugements de valeur sur la base des faits ou des conséquences du jugement ;
- la **clarification avancée** : utilisation de définitions, clarification à partir d'indices pragmatiques même à propos de contenus qui ne sont pas explicitement exprimés, utilisation des assomptions ;
- les **suppositions** ;
- la **synthèse**.

Ces capacités sont listées comme si elles étaient applicables dans n'importe quel contexte, tout comme les dispositions. Cependant, Ennis souligne à la fois l'importance de s'adapter au contexte et celle des connaissances de domaine en association avec les capacités indiquées. Enfin, Ennis donne une place fondamentale aux dispositions pour l'EC. L'ensemble des dispositions dessine donc le penseur critique idéal comme étant, à la fois, à la recherche d'explications alternatives, ouvert, prêt à prendre au sérieux les points de vue des autres, prêt à changer d'idée face à de nouvelles justifications, mais aussi prêt à se donner du mal pour repérer l'information pertinente, pour chercher à comprendre et pour chercher à clarifier son point de vue face aux autres, comme un agent compréhensif, conscient du contexte.
L'éducation de l'EC implique donc de viser ces capacités et dispositions. L'évaluation de l'EC consiste quant à elle à examiner l'utilisation de ces mêmes compétences et la présence des dispositions indiquées (Ennis est à l'origine de tests standardisés, encore largement utilisés : les tests de pensée critique de Cornell niveaux X et Z et le test de pensée critique Ennis-Weir).

Presque en parallèle avec le travail de Robert Ennis se développent le programme *Philosophy for Children* et le *Center for Critical Thinking*.

Matthew Lipman et le programme *Philosophy for Children*

Le programme *Philosophy for Children* (P4C) naît dans les années 1970 aux États-Unis, à l'initiative de Matthew Lipman (voir, par exemple, Lipman 1987, 2003). Parmi ses objectifs, ce programme inclut l'éducation de l'esprit critique. Lipman considère que les enfants sont capables de mobiliser l'EC en particulier lorsqu'ils sont stimulés dans cette direction par le programme P4C, et qu'ils devraient être encouragés à le faire dès leur entrée à l'école. Son approche consiste à identifier l'EC avec le concept de jugement. Le jugement est ce qui permet de former des opinions, des évaluations, des conclusions. L'EC permet donc de juger correctement[13]. Lipman oppose ainsi le jugement ordinaire, les jugements que nous émettons tous au quotidien, au jugement critérié, en quelque sorte spécialisé, avancé, et connecté à l'EC (*"making judgment with criteria" = "good judgment"* et *"making judgment without criteria"*

= "*ordinary judgment*"). L'EC doit être éduqué parce que c'est ainsi que nous allons rendre possible l'exercice de notre jugement de manière avancée, « professionnelle[14] ». Comme l'architecture est la manière « pro » de regarder un bâtiment, l'EC est alors la manière avancée d'exercer ses capacités de jugement.

"The outcomes of critical thinking are judgments; and the nature of judgment is such that critical thinking may be defined as skillful, responsible thinking that facilitates good judgment because (1) it relies upon criteria; (2) it is self-correcting, and (3) it is sensitive to context. The very meaning of 'criterion' is 'a rule or principle utilized in the making of judgments.' Judgment, in turn, is a skill; therefore critical thinking is skillful thinking, and skills can only be defined through criteria by which performance can be evaluated. So critical thinking is thinking that both employs criteria and can be assessed by appeal to criteria." (Lipman 1988[15])

Éduquer l'EC signifie donc donner accès à ces critères de jugement avancé. Les notions de jugements et de critères sont ainsi interdépendantes : les critères sont des règles ou des principes utilisés pour émettre un bon jugement, des raisons particulièrement fiables. Ils sont spécifiques à un certain domaine, d'où l'importance de prêter attention au contexte dans le choix des critères appropriés et de la connaissance des contenus liés au contexte lui-même. Cependant, certains critères comme la cohérence, la fiabilité et d'autres sont des méta-critères parce qu'ils ont un tel niveau de généralité qu'ils s'avèrent toujours nécessaires pour l'EC. Dans des domaines spécifiques, les critères sont normalement très bien acceptés et respectés par les experts de la communauté d'enquête. L'utilisation compétente de ces raisons respectées est un moyen d'établir l'objectivité de nos jugements normatifs, descriptifs et évaluatifs. En plus des critères, l'EC comprend une composante métacognitive avancée, qui ne se limite pas au fait de se regarder, de se monitorer, mais qui implique également l'autocorrection.
Lipman voit aussi l'éducation de l'EC comme une formation de la responsabilité intellectuelle. En montrant aux élèves des modèles de responsabilité épistémique, les enseignants les invitent à assumer la responsabilité de leur propre pensée et, dans un sens plus large, de leur propre éducation.

La position de Lipman est proche d'autres approches philosophiques en ce qui concerne l'appel au jugement, aux critères, à la métacognition, et naturellement la nature normative et volontaire de l'exercice de l'EC. Mais elle présente aussi des spécificités. Tout d'abord, sa définition est restreinte et opérationnelle : selon lui, l'EC consiste à mettre en place des critères et une certaine attitude métacognitive. Les critères peuvent être identifiés et listés – même si, en grand partie, ils présentent une spécificité de domaine. Ensuite, Lipman met en évidence la nature « experte » de l'EC, qui permettrait de passer d'un niveau de jugement ordinaire à un niveau avancé, expert, rassemblant à celui des professionnels. L'EC serait donc une forme de pensée outillée (par des critères), qui se corrigerait, prendrait en compte le contexte et, pour cette raison, serait une forme de pensée experte. Lipman ne se pose cependant pas le problème de la formation au quotidien d'un jugement non expert : quelles bases cognitives naturelles le permettraient ? Quels critères implicites seraient employés ? Quelles contraintes limiteraient ce type de jugement ? etc.

Le Center for Critical Thinking et sa fondation

Dans les années 1980 naissent le *Center for Critical Thinking* et la *Foundation for Critical Thinking*, à l'initiative de Richard Paul, philosophe, Linda Elder, enseignante, et Gerald Nosich, philosophe (*Foundation for Critical Thinking*, https://www.criticalthinking.org/ ; (Lewis & Smith 1993[16]). Richard Paul est aussi à l'origine d'articles, d'un test d'EC et de plusieurs livres contenant des indications méthodologiques pour son enseignement (voir : Scriven & Paul 1987, Paul & Elder 2009, Elder & Paul 2010). Son approche part du double constat de la nature universelle de la pensée et de la qualité universellement insatisfaisante des

produits de la pensée. En effet, la pensée humaine est, selon lui, influencée par des préjugés, des mythologies, des illusions, de l'ignorance, et elle a tendance à s'auto-tromper. Les capacités à cultiver sont de ce fait liées aux différents processus composant la pensée. Paul en identifie huit :

- générer des buts ;
- soulever des questions ;
- utiliser l'information ;
- utiliser des concepts ;
- faire des inférences ;
- faire des assomptions ;
- générer des implications ;
- prendre un point de vue.

Le penseur critique met en place ces processus en utilisant certaines capacités, comme le penseur non critique, mais, contrairement à ce dernier, le penseur critique adopte lui des standards intellectuels exigeants pour chaque capacité.

Paul met surtout l'accent sur l'importance des traits de personnalité ou dispositions à côté des capacités : ces dernières ne sont pas suffisantes à produire une pensée critique au sens fort (« *strong-sense critical thinking* ») mais seulement une pensée critique au sens faible (« *weak-sense critical thinking* »). Ces traits de personnalité sont aussi appelés des « vertus épistémiques », qu'il faut donc cultiver de pair avec les capacités propres à l'EC pour devenir un penseur critique au sens fort. En effet, selon Paul, cultiver les seules capacités de l'EC peut conduire aussi bien à « faire le bien » qu'à son contraire, car rien dans les capacités de pensée ne garantit qu'elles amènent à des choix éthiques.

Pour que l'EC soit utilisé à bon escient, à des fins éthiques, il est donc nécessaire de développer les vertus épistémiques de l'empathie intellectuelle (qui permet de comprendre le point de vue des autres), de l'intégrité, de la persévérance, du courage, de l'autonomie, de l'amour de la vérité et de la raison, de la justice, et de la « *fair-mindedness* ». L'ouverture d'esprit et l'honnêteté intellectuelle (*fair-mindedness* en anglais) est associée à l'EC au sens fort. Paul distingue ainsi un EC « socratique » (qui exerce ces vertus en plus des capacités) et un EC « sophiste » (égoïste, non éthique), utilisé dans le but de tromper les autres, de les persuader pour en abuser. Paul distingue également EC « explicite » et EC « implicite » (lorsque les capacités sont utilisées habilement mais sans prise de conscience de leur rôle et de la manière dont elles peuvent être exercées de manière améliorée), EC « systématique » et EC « épisodique » (ce dernier étant exercé seulement sur des contenus particuliers), EC « émancipateur » et EC « contraint » (qui utilise des procédures et stratégies mais n'est pas ouvert à considérer d'autres points de vue et alternatives), EC basé sur des langages « naturels » ou « formels » (Elder & Paul 2010). Au total, Paul & Elder (2009) listent 35 « dimensions » de l'EC qui comprennent des dimensions affectives (relatives aux vertus intellectuelles), des macro-capacités et des micro-capacités.

La définition de l'EC est celle d'une forme de pensée disciplinée, qui utilise activement différentes capacités intellectuelles pour guider la formation de croyances et l'action. Elle est universelle et générale :

"*Critical thinking is the intellectually disciplined process of actively and skillfully conceptualizing, applying, analyzing, synthesizing, and/or evaluating information gathered from, or generated by, observation, experience, reflection, reasoning, or communication, as a guide to belief and action. In its exemplary form, it is based on universal intellectual values that transcend subject matter divisions: clarity, accuracy, precision, consistency, relevance, sound evidence, good reasons, depth, breadth, and fairness*." (Scriven & Paul 1987)

Du point de vue pratique, l'enseignement de l'EC est, pour Paul, explicite et focalisé : il est nécessaire d'insérer dans les curricula les compétences de l'EC et de garantir qu'elles soient visées dans toutes les matières scolaires. Un apprentissage approfondi des compétences de l'EC demande une longue pratique et instruction mais tout le monde peut réussir. Une condition essentielle pour que l'enseignement de l'EC réussisse est que les enseignants soient eux-mêmes capables et disposés à penser de manière critique, et qu'ils aient compris non seulement quelles compétences ils doivent cultiver, mais aussi les fondements de l'EC. De plus, il est important qu'ils enseignent les compétences de l'EC, mais également qu'ils cherchent à développer les dispositions pour l'EC (Elder & Paul 2010).

Le rapport Delphi

L'intérêt croissant pour l'éducation de l'esprit critique fait que, entre 1987 et 1989, un groupe de 46 philosophes et éducateurs se réunit en un panel Delphi afin d'atteindre un consensus concernant la définition de l'EC, les capacités et les dispositions impliquées, et pour définir des modalités de mesure de l'EC. Le panel est coordonné par Peter Facione et commissionné par la *American Philosophical Association* (Facione 1990). Le résultat constitue une définition normative de l'EC, assez large pour mettre les participants d'accord.

"CONSENSUS STATEMENT REGARDING CRITICAL THINKING AND THE IDEAL CRITICAL THINKER"
We understand critical thinking to be purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based. CT is essential as a tool of inquiry. As such, CT is a liberating force in education and a powerful resource in one's personal and civic life. While not synonymous with good thinking, CT is a pervasive and self-rectifying human phenomenon. The ideal critical thinker is habitually inquisitive, well-informed, trustful of reason, open-minded, flexible, fair-minded in evaluation, honest in facing personal biases, prudent in making judgments, willing to reconsider, clear about issues, orderly in complex matters, diligent in seeking relevant information, reasonable in the selection of criteria, focused in inquiry, and persistent in seeking results which are as precise as the subject and the circumstances of inquiry permit. Thus, educating good critical thinkers means working toward this ideal. It combines developing CT skills with nurturing those dispositions which consistently yield useful insights and which are the basis of a rational and democratic society." (Facione 1990[17])

Facione reconnaît la nécessité et la difficulté de parvenir à une définition satisfaisante d'un point de vue conceptuel qui ne soit ni trop restreinte, ni trop large. Il invite à identifier des exemples de « bons » et de « mauvais » penseurs critiques et, de là, à en extraire les capacités et attitudes nécessaires pour définir l'EC. Il soutient que ce type de procédé a influencé le travail du groupe d'experts réunis dans le panel Delphi (Facione 2011). À la fin de ce travail, les experts du rapport Delphi ont ainsi identifié six capacités fondamentales dans lesquelles le bon penseur critique excellerait mais que le mauvais penseur critique aurait du mal à mener à bien : interpréter, analyser, évaluer, faire des inférences, expliquer et s'autoréguler (Facione 2011[18]). Cependant, ces capacités sont idéales.

"*The experts articulated an ideal. It may be that no person is fully adept at all the skills and sub-skills the experts found to be central to CT. It may be that no person has fully cultivated all the affective dispositions which characterize a good critical thinker. It may be that humans compartmentalize their lives in ways that CT is more active and evident in some areas than in others. This gives no more reason to abandon the effort to infuse CT into the educational system than that knowing no friendship is perfect gives one reason to despair of having friends. The experts' purpose in putting the ideal before the education community is that it*

should serve as a rich and worthy goal guiding CT assessment and curriculum development at all educational levels." (Facione 1990, Facione 2011[19]).

Sur la question de la nécessité, pour le « bon » penseur critique, de faire également preuve de vertus morales (utiliser l'EC pour le bien), le panel Delphi ne propose pas de position consensuelle : une minorité considère en effet que l'EC ne peut pas être disjoint de l'idée d'éthique alors que la majorité considère, elle, l'EC comme un ensemble de dispositions et de capacités ne comportant pas de positionnement moral (ou politique, religieux, etc.).

Peter Facione

Peter Facione a par la suite développé des outils de mesure de l'EC, comme le *California Critical Thinking Test* (CCTST) qui s'inspire directement de la définition ayant émergée du travail des experts réunis pour le rapport Delphi. Facione a également mis au point des tests spécifiquement adressés aux dispositions de l'EC (CCTDI). Comme les autres philosophes s'étant intéressés à l'EC, Facione s'est demandé si posséder des capacités suffisait à faire preuve d'EC. Il considère qu'il est improbable de posséder des capacités à un niveau avancé et de ne pas les utiliser. Cependant, la possibilité de ne pas utiliser ses capacités existe, et il faut donc aussi fournir une liste de dispositions faisant du penseur critique non seulement un penseur habile, mais aussi une sorte de Sherlock Holmes motivé par la recherche de la vérité. (L'EC est en effet, selon Facione, « l'ensemble d'attitudes et capacités qui nous donne les meilleurs chances d'arriver à la vérité ».)
Ainsi, un penseur critique possède selon Facione (et conformément à la vision du panel Delphi) une certaine attitude envers la vie – active, rationnelle, ouverte sur le point de vue des autres, honnête… – et fait preuve de certaines attitudes lorsqu'il s'agit de faire face à un problème, une question : clarté, systématicité, discipline, organisation, attention, précision. Inversement, un faible penseur critique se comporte de manière désorganisée et simpliste lorsqu'il s'agit de résoudre un problème, il ne cherche les informations nécessaires que de manière ponctuelle et non systématique, il se laisse facilement distraire et il est prêt à tout laisser tomber face à la première difficulté, se satisfaisant d'une réponse vague et générale (Facione 2011).

Facione se réfère à des travaux en sciences cognitives concernant le raisonnement et notamment l'articulation de la pensée en deux systèmes complémentaires : intuitif et réflexif. Il est très attentif à ne pas les mettre en opposition, et il les considère comme deux systèmes opérant en parallèle (et non de façon séquentielle). Il fait ainsi référence aux notions de Système 1 et Système 2, dont il fournit une vision simplifiée. (En effet, les concepts de processus de pensée de type 1 et de type 2 font l'objet de plusieurs définitions, et sont soumises à débat au sein de la communauté scientifique.) Enseigner l'EC permet de développer les processus propres au Système 2, mais les deux systèmes sont nécessaires pour la bonne prise de décision. Ainsi, les erreurs dans la prise de décision et les erreurs de jugement peuvent naître – selon ce qu'affirme Facione – aussi bien d'erreurs du Système 1 (des heuristiques mal appliquées ou qui influencent à tort un jugement) que d'erreurs du Système 2 (de la mauvaise représentation des faits, etc.). Facione poursuit en décrivant un certain nombre d'heuristiques dont l'activité inappropriée est susceptible de nous amener à l'erreur : disponibilité, affect, association, simulation, similitude, satisfaction, aversion au risque et à la perte, ancrage, illusion de contrôle, biais rétrospectif, dominance (Facione, 2011[20]).

D'un point de vue pratique, afin d'aider l'aspirant penseur critique à développer ou à mettre en place ces fonctions, Facione propose une liste de questions à se poser face à une situation qui nécessiterait de l'EC (cette *check-list* fait partie du manuel CCTST de test des capacités de l'EC) et une liste à points pour l'auto-évaluation de la disponibilité à faire preuve de EC – ou du moins des dispositions dont le sujet aurait fait preuve dernièrement (cette liste est une version simplifiée du CCTDI ou des test des dispositions de l'EC).

Sharon Bailin : l'enquête critique

Sharon Bailin et ses collègues, déplorent le fait que la notion d'EC, bien que répandue notamment dans les curricula et les discussions autour des curricula aux États-Unis et au Royaume-Uni, demeure vague et soit l'objet d'un consensus seulement superficiel, qui existe tant qu'on ne touche pas aux questions pratiques sur la manière de l'enseigner, de caractériser le penseur critique, des objectifs à atteindre. Ces auteurs reconnaissent que le flou autour de la notion représente un problème pour les éducateurs (Bailin et al. 1999).

Bailin, philosophe (professeure à la *Faculty of Education* de la Simon Fraser University), est la représentante d'une approche canadienne de l'EC qui se démarque sous plusieurs aspects de l'états-unienne. Son approche est fondamentalement disciplinaire : le but de l'éducation à l'EC étant de « bien penser » au sein d'une discipline. De plus, cette approche est fondée sur l'investigation. En réalité, le terme « investigation » est utilisé par Bailin et ses collaborateurs (comme Mark Battersby, du département de philosophie de Capilano University) dans un sens large, et inclut l'investigation menée de manière argumentative, en quête des meilleurs arguments et des meilleures raisons pour soutenir une prise de position ou une opinion. L'enseignement de l'EC est ainsi intégré au niveau des disciplines et sert à apprendre à utiliser les normes et les modes de raisonnement propres à la discipline. Afin de penser de manière critique et d'apprendre à utiliser normes et modes de raisonnement, apprendre les concepts de la discipline n'est pas suffisant. C'est pour cette raison que l'enseignement disciplinaire se doit d'intégrer une approche bien précise de l'EC en son sein : raisonnement et argumentation doivent faire l'objet d'une attention explicite et focalisée dans chaque enseignement disciplinaire. Pour cela, l'enseignement disciplinaire doit intégrer, d'une part, des aspects d'argumentation propres à la discipline et, d'autre part, des éléments concernant l'argumentation qui sont d'ordre plus général, transdisciplinaire. Par exemple, il est important d'apprendre à connaître les différents arguments faisant débat au sein de la discipline, les débats historiques, etc. Sans cette approche disciplinaire, il n'est pas possible d'apprendre à bien raisonner et à bien argumenter au sein de la discipline elle-même. Par exemple, un cours d'écologie qui intègrerait la dimension d'EC en tant qu'investigation amènerait les étudiants à cerner le contexte et les différents arguments pour et contre une certaine décision, et les pousserait à se demander quelles sont les règles permettant d'évaluer correctement le poids de chaque argument et de ses raisons.

Bailin reconnaît son affinité avec les approches de différents philosophes et éducateurs comme Richard Paul, Robert Ennis et Matthew Lipman. Ainsi, l'EC est selon elle une forme de pensée qui est dirigée vers un objectif (contrairement, par exemple, à la rêverie) tel que prendre une décision, répondre à une question, résoudre un problème. Plus généralement, il s'agit d'une forme de pensée permettant de former un jugement, de développer une idée sur ce qu'il faut croire, ou non. Bien évidemment, il ne s'agit pas de n'importe quel type de pensée qui amène à prendre une décision.
L'EC comporte une dimension normative : une décision prise sur la base de raisons pauvres ou d'un processus de raisonnement faible ne compterait pas comme de l'EC. L'EC doit donc se conformer à des standards ou critères permettant de déterminer si un raisonnement est de bonne ou de mauvaise qualité. En outre, pour affirmer faire preuve d'EC, ces critères et standards doivent être utilisés de manière consciente et volontaire, non de façon aléatoire, même si la personne qui les utilise pourrait ne pas être capable de les formuler verbalement. Ce type de définition se conforme, selon Bailin, à l'idée que les éducateurs se font du concept d'EC (Bailin et al. 1999[23]).

Cependant, ce type de définition présente des défauts majeurs, notamment en termes opérationnels. Tout d'abord, elle est vague : ni les jugements ni les standards ne sont définis. Ensuite, les limites de l'EC par rapport à d'autres formes de pensée ne sont pas claires (par exemple, quelle différence avec la résolution de problèmes, qui demande de former des

jugements pour arriver à la décision ?). De plus, le concept peut être à la fois trop large et trop restreint. Bailin cite la définition de Seigel de 1988 qui limite l'EC à la capacité d'évaluer correctement les raisons et à la disposition et volonté de le faire pour orienter le choix de ses croyances et actions. Bailin considère cependant cette vision de l'EC comme trop restreinte : évaluer la qualité des arguments ne suffit pas ; il faut également savoir quelles raisons sont pertinentes, savoir en prendre en compte d'autres, savoir voir les alternatives. Enfin, ce type de définition ne dit pas comment on arriverait à atteindre l'objectif. Mais, surtout, pour Bailin, il ne prend pas en compte le fait que, le plus souvent, l'exercice de l'EC a lieu dans un contexte social, dans le cadre d'une discussion avec un échange d'arguments. C'est pour cette raison qu'elle propose une approche basée sur la discussion critique – ce qu'elle appelle l'investigation, qui est aussi un élément fondamental de la production de connaissances scientifiques (selon Popper, que Bailin cite). Le déploiement de l'EC ne se limiterait donc pas à l'évaluation des raisons et des arguments, mais comporterait un exercice permettant de répondre correctement aux arguments d'autrui dans le contexte d'une discussion (Bailin et al. 1999[22]).

Une fois la dimension sociale et argumentative de l'EC établie, il est nécessaire de définir les standards auxquels l'évaluation des raisons et l'échange d'arguments doivent se conformer afin de constituer une forme d'EC – rappelons que Bailin a une approche philosophique, normative : ce qui compte est d'identifier, établir, les standards qui mènent à bien penser et non de décrire les capacités permettant de penser. Faire preuve d'EC consisterait donc à appliquer les normes qui rendent correct un raisonnement. Les capacités à cultiver pour éduquer l'EC sont celles qui permettent de se conformer aux standards pertinents[23].

À la différence des auteurs de la tradition états-unienne, Bailin refuse d'utiliser les termes « capacités » ou « habiletés » pour décrire ce que le penseur critique doit faire. Elle réfute donc le jargon apparemment descriptif et psychologique des autres philosophes. En effet, selon elle, ce que ces philosophes indiquent est en réalité une liste d'objectifs à atteindre, de tâches que le penseur critique doit être capable de réaliser, et non des habiletés qu'il possèderait. Bailin résout donc l'ambiguïté en choisissant d'appeler ces « capacités » des « résultats » que le penseur critique doit viser et obtenir (Bailin et al. 1999[24]).

Bailin est complètement agnostique quant aux réelles capacités cognitives qu'il est nécessaire d'utiliser et de développer afin d'atteindre ces objectifs. De plus, elle considère qu'enseigner l'EC ne signifie pas faire apprendre un ensemble de capacités, mais faire acquérir un ensemble de ressources. Ces ressources sont de cinq types (Bailin et al. 1999[25]) :

- des connaissances factuelles, spécifiques à un domaine (*background knowledge*). La capacité de penser dans un domaine de manière claire, précise, approfondie dépend essentiellement des connaissances que l'on possède dans ce domaine. La conception de l'EC portée par Bailin est ainsi fortement contextualisée et dépendante du contexte et du contenu, contrairement à d'autres conceptions plus générales de l'EC. En effet, Bailin fait dépendre l'EC des standards qui sont appliqués et du fait d'utiliser ces standards et non de soi-disant capacités ou habiletés cognitives. Même des capacités apparemment générales comme celle de réaliser une inférence correcte pour une généralisation dépendent du fait de posséder un minimum de connaissances à propos du domaine du concept à généraliser ;
- des standards d'évaluation critique. Ces standards, à leur tour, sont spécifiques à un domaine. Par exemple, la science possède des standards et normes qu'elle applique pour décider quelle est la meilleure hypothèse, explication, etc. Pour penser de manière critique en science, il faut donc connaître et savoir appliquer ces standards, par exemple les critères qui font d'une donnée une preuve de bonne ou moins bonne qualité. Cependant, les standards ne sont pas les mêmes en sciences et en art ou dans le cadre des choix éthiques. Afin de penser de manière critique dans chacun de ces domaines, il faut donc acquérir les standards requis. Il ne s'agit pas d'en fournir une liste, ou de les apprendre à titre d'éléments théoriques : afin d'être capable de les

utiliser de manière appropriée, il est nécessaire de comprendre les pratiques dans lesquelles ces standards opèrent. Pour appliquer les standards de la science de manière appropriée, par exemple, il ne suffit pas de connaître les principes de la méthode expérimentale, mais il faut posséder une connaissance de la science en tant que pratique. Les standards des différentes disciplines ou domaines sont d'ailleurs des produits culturels, qui évoluent et peuvent être soumis à critique ;

- comme chaque standard est propre à un domaine, le penseur critique doit être capable de reconnaître et de distinguer différents domaines pour choisir les standards de manière appropriée (par exemple, distinguer entre domaine moral et domaine scientifique) ;
- le penseur critique possède aussi des heuristiques. Ce terme ne désigne pas les heuristiques cognitives naturelles, mais les stratégies historiquement inventées pour mieux comprendre et analyser des concepts, etc. Par exemple, le fait d'avoir recours à des contre-exemples est une heuristique qui aide à clarifier les aspects d'un concept ;
- le penseur critique a des habitudes de pensée, des attitudes, qui le mettent en condition d'utiliser les autres ressources. Cependant, ces habitudes ne sont pas uniquement des automatismes ; le penseur critique dispose d'un esprit critique : il reconnaît la valeur de l'EC, son importance et, pour cette raison, il choisit d'utiliser les ressources nécessaires. En cela, Bailin s'accorde avec la tradition états-unienne, qui met l'accent sur les dispositions, et elle cite à cet effet les dispositions classiquement évoquées par Paul, Ennis ou le rapport Delphi, telles que le respect de la raison et de la vérité, le questionnement, la volonté d'investiguer, l'ouverture d'esprit, l'indépendance d'esprit, la volonté d'être juste (*fair-mindedness*), le respect des autres, le respect des autorités intellectuelles légitimes, l'éthique du travail bien fait.

L'éducation de l'esprit critique vise à favoriser l'utilisation des concepts et standards que notre culture a développé pour améliorer la pensée et la rendre plus fructueuse. Le moyen principal pour atteindre cet objectif consiste, selon Bailin, à se familiariser, par le biais d'exemples, avec l'utilisation de ces principes, et à comprendre quand ils s'appliquent et comment. Bailin prône donc un enseignement contextualisé, dans des contextes disciplinaires riches, complexes et réalistes. Cette modalité d'enseignement s'oppose à celle consistant à présenter des règles de pensée et des principes de manière abstraite. L'enseignement de l'EC doit commencer avant l'école et se poursuivre sur des questions de plus en plus complexes et spécialisées, et en mettant de plus en plus l'accent sur l'intérêt d'utiliser ces principes, donc sur une utilisation consciente et motivée de l'EC. Il s'agit donc d'un processus graduel d'apprentissage (Bailin et al. 1999[26]).

Le processus permettant de parvenir à ces jugements raisonnés est celui de l'« *inquiry* » ou « *critical inquiry* ». Dans ses manuels d'enseignement de l'esprit critique, Bailin présente les indications nécessaires afin de bien mener cette méthode (voir, par exemple, Bailin & Battersby 2016). Ce processus comprend différents aspects : identifier la question de manière spécifique (*identifying the issue*), identifier le contexte (*identifying relevant contexts*), comprendre les positions opposées (*understanding the competing cases*), comparer et baser ses jugements sur ces comparaisons (*making a comparative judgment among them*). Cette approche théorique motive un choix pédagogique consistant à mettre les apprenants dans une vraie situation dialogique, c'est-à-dire leur présenter une discussion autour d'une question débattue, en fournissant le contexte et l'historique du débat. Dans la discussion sont présentés différents arguments en faveur ou contre une certaine prise de position. Les dialogues peuvent porter sur des arguments aussi divers que le végétarianisme, la peine de mort, la théorie de l'évolution, la violence dans les média, la nature humaine, la polygamie, l'interprétation des œuvres d'art, les théories du complot… Ce type d'intervention peut donc être inséré dans n'importe quelle matière scolaire.

La méthode pédagogique présente une théorie de l'argumentation : une explication de ce qui constitue un « bon argument », la présentation des supercheries argumentatives, de la

déduction et l'induction, des arguments analogiques, etc. De plus, elle comprend une méthodologie pour l'analyse des sources (sources internet, citations, etc.). Ces éléments constituent des heuristiques à mobiliser pour rendre la pensée plus efficace.
La méthode pédagogique se focalise donc sur la manière de bien conduire une investigation, c'est-à-dire de bien identifier la question qui est débattue, de bien différencier le type de jugement qu'on veut donner (évaluation ou plutôt jugement factuel). Les apprenants apprennent ensuite comment préparer leur comparaison des arguments présentés : ils doivent d'abord les identifier un par un comme étant pour ou contre la position débattue, identifier les raisons apportées, etc. Ensuite, ils doivent bien comprendre le contexte historique et social de la question débattue. Enfin, ils doivent analyser la force de chaque argument apporté, en utilisant des critères différents. Si un argument est de type éthique, les critères sont d'une autre nature que s'il est de type factuel. C'est à ce moment qu'il est possible de se forger une opinion à partir de la comparaison des arguments apportés. Cependant, le résultat n'est pas garanti comme s'il existait un algorithme permettant de calculer les pours et les contres de manière objective. Bailin rappelle qu'il est important de mesurer ses jugements par rapport à son niveau de confiance.

Finalement, la méthode pédagogique proposée propose des stratégies pour dépasser les obstacles rendant difficile de garder son esprit ouvert face aux autres positions et arguments ; Bailin reconnaît en effet que la simple exhortation à rester ouvert n'est pas suffisante pour maintenir la « bonne attitude ». Puisque l'approche de l'EC proposée par Bailin est dialectique et argumentative, chaque chapitre de son manuel (Bailin & Battersby 2016) commence par une narration contrastant de discussions critiques et non critiques sur un sujet. Les lecteurs sont invités à faire une pause et à essayer de développer leur propre avis sur les questions présentées (Bailin & Battersby 2016[27]).

L'approche pédagogique de Bailin répond à un constat qu'elle a elle-même soulevé contre les méthodes « générales » d'éducation de l'EC : la nature de la pensée est contextualisée. Et ceci non seulement pour des raisons cognitives – le fait que la pensée est toujours celle d'un objet dans un contexte – mais aussi pour des considérations d'ordre culturel : nous avons, dans le cadre de notre culture, des normes et des standards qui nous permettent de penser de manière plus experte, claire, précise, approfondie, et ceci dans le cadre d'une variété de disciplines et avec un degré différent de développement. Si nous voulons apprendre à mieux penser, nous ne pouvons pas nous détacher de ce contexte culturel. Se pose alors le problème du transfert. Si les entraînements des soi-disant capacités, notamment dans le cadre des approches plus généralistes, ont pour Bailin peu de chances de produire des résultats, son approche risque tout de même de rendre l'EC tellement spécialiste qu'il n'y aurait aucun espoir de le développer dans un domaine puis de le transférer dans un autre. Bailin est conscience de ce problème. Sans apprendre les règles d'une certaine discipline ou d'un certain domaine, nous ne pouvons pas espérer réussir à penser de manière réellement plus experte dans le domaine en question. Cependant, Bailin ne soutient pas qu'il n'existe pas de place pour une éducation de l'EC qui se situerait en amont des domaines disciplinaires (en matière de curricula scolaires). Au contraire, elle préconise un enseignement progressif qui commencerait avant l'entrée à l'école. Elle propose également de baser l'enseignement de l'EC sur des exemples concrets.

Sa proposition consiste donc à fournir des exemples variés, d'introduire l'enseignement de l'EC dans toutes les disciplines pour que l'élève apprenne une variété de standards et normes qui s'appliquent à plusieurs domaines. L'élève deviendra ainsi plus performant dans le type de jugement attendu dans différentes disciplines : sciences physiques, sciences sociales, arts, etc. – c'est-à-dire dans une variété de contextes dans lesquels il est pour lui important d'apprendre à bien penser. Il existe cependant des problèmes qui tombent en dehors des disciplines scolaires classiques, par exemple des problèmes d'ordre moral ou de la vie quotidienne – comme le montre l'exemple du végétarianisme – mais même au-delà : comment choisir son université ? Pour qui voter aux prochaines élections ? Bailin reconnaît qu'il est important de développer ses habitudes de pensée et d'acquérir des standards, des

connaissances, des heuristiques également pour mieux considérer ces questions. Elle se dit donc favorable à ce que l'EC soit aussi enseigné dans des cours spécifiquement dédiés (et non seulement par infusion dans les diverses disciplines), pourvu qu'il soit enseigné dans un contexte dense, que les ressources à acquérir soient fournies à partir d'exemples concrets et riches de contenu et non comme des listes de standards, critères ou prétendues capacités à développer.

Caractères communs aux approches philosophiques

Les approches philosophiques partagent certains points (Bailin 1999, Siegel 2010[28]) :

- leur nature est normative : les définitions philosophiques indiquent les objectifs à atteindre et les standards à respecter. D'un point de vue théorique comme pratique, il s'agit d'identifier des normes, des standards, des critères à appliquer pour atteindre un objectif cognitif ou pratique : rendre la pensée plus efficace pour atteindre un but désigné, augmenter les chances de prendre une meilleure décision, embrasser une opinion correcte ou proche de la réalité, éviter de développer des croyances fausses et/ou non justifiées ;
- ces approches sont basées sur la polarisation entre penser (bien penser) et ne pas penser : penser est volontaire, explicite, et est le contraire de l'action immédiate. Elles font souvent référence à l'idée de « penseur idéal » : rationnel, motivé à chercher la vérité, doté des capacités pour le faire ;
- leur nature est explicite : ces définitions font uniquement référence à des capacités utilisées de manière volontaire et assumée, réflexive ;
- ces approches fournissent le plus souvent une liste des capacités attribuées à la pensée ou esprit critique : métacognition au sens très large, jugement, rationalité ou raison et recours aux raisons (donner et évaluer des raisons ou arguments), inférence, inférence logique, capacités de communication, expression, clarté, etc. Cette liste est souvent longue (même quand un nombre restreint de catégories est identifié, celles-ci sont alors divisées en sous-catégories). La liste est aussi très variée, et souvent non spécifique à l'EC. Par exemple, elle comprend des capacités comme celle d'interprétation de textes, ou celle de réaliser des inférences, la capacité de juger, celle d'analyser des arguments ou de s'exprimer correctement ;
- point critique : les capacités indiquées manquent de réalisme cognitif. Même si la littérature philosophique parle souvent de capacités (« *skills* », « *abilities* »), en réalité, comme le fait remarquer Sharon Bailin, ces capacités ne sont pas décrites dans leur fonctionnement « naturel », mais plutôt identifiées comme des buts à atteindre. Il s'agit, par exemple, de bien évaluer l'information ou de former des inférences correctes. Or cette littérature ne décrit pas comment les inférences sont formées et donc comment les améliorer, sauf en indiquant quels critères adopter lorsque nous formons une inférence pour qu'elle soit correcte. Le penseur critique est un penseur idéal qui n'a pas besoin d'exister dans la réalité, car ses caractéristiques sont des perfections de la pensée. Souvent, les approches philosophiques font référence aux obstacles représentés par la pensée naturelle, mais elles ne se demandent pas si les perfections de pensée proposées (capacités mais aussi dispositions) sont réalistes : dans quelles circonstances peuvent-elles être réalisées ? dans quelles limites ?
- les approches philosophiques font référence, en plus des capacités, à des dispositions, voire des vertus épistémiques. Cependant, la question des dispositions et des vertus reste encore aujourd'hui ouverte dans le débat philosophique autour de l'EC. Harvey Siegel met un accent particulièrement fort sur le fait que la définition philosophique de l'EC dispose de deux composantes : l'une, cognitive, liée aux capacités et l'autre, émotionnelle, relative aux dispositions [il appelle « *reasoned assessment* » (évaluation raisonnée) la première et « *critical spirit* » (esprit critique) la seconde], mais que les

deux ne devraient pas être considérées séparément. Ainsi, l'idée d'un penseur critique idéal prend une dimension fortement personnelle, et devient encore plus « idéalisée » (Siegel 2010[29]) ;

- les représentants des différentes approches philosophiques proposent souvent des manuels pédagogiques de l'EC et des tests qui tendent à vérifier si les critères, normes et standards sont correctement appliqués, de manière disciplinaire ou a-disciplinaire, dans des contextes plus ou moins réalistes, avec un transfert plus ou moins éloigné des situations d'apprentissage. Certains préconisent un enseignement plutôt général de l'EC, d'autres un enseignement infusé dans les disciplines. Ceci dépend de la vision qu'ils ont de l'EC en tant que capacité générale ou spécifique. Plusieurs chercheurs considèrent les connaissances de domaine (*background knowledge*) comme nécessaires pour bien penser à l'intérieur de ce même domaine (par exemple, Bailin 2002 ; Willingham 2007). Pour d'autres, des critères d'ordre général existent et sont transversaux aux domaines de connaissance, même si ceux-ci possèdent des critères propres (par exemple, Lipman 1987 ; McPeck 1990). (Pour une discussion à ce propos, voir Ennis 1989, Lai 2011[30].)

2.1.2 Les approches psychologiques

Aux approches philosophiques s'ajoutent, dans le domaine de la réflexion sur l'EC, un certain nombre d'approches psychologiques. À l'origine de ces approches, nous trouvons un groupe varié de psychologues tels que Diane Halpern (psychologie cognitive), Deanna Kuhn (psychologie de l'éducation, psychologie du développement), Richard Nisbett (psychologie sociale, culture et cognition), Keith Stanovich (psychologie cognitive, psychologie du raisonnement), Robert Sternberg (psychologie, psychométrie), Tim Van Gelder (philosophie), Daniel Willingham (psychologie de l'éducation). Diane Halpern est la seule de ce groupe à avoir proposé un manuel assez complet pour l'enseignement de l'EC aussi bien qu'un test commercialisé (Halpern Critical Thinking Assessment - HTCA). Richard Nisbett a mené des études sur l'efficacité de méthodes d'enseignement adoptées par ses collègues et lui auprès d'étudiants de niveau universitaire et des adultes. Keith Stanovich a développé une théorie du fonctionnement de l'esprit – la théorie duale ou des « deux systèmes » – qui est souvent citée par d'autres approches psychologiques mais également philosophiques (Facione) comme base de l'EC. Deanna Kuhn se distingue car elle a proposé la seule approche développementale et réellement descriptive de l'EC : elle a en effet cherché à identifier les bases cognitives qui constitueraient le background naturel de l'EC avancé, éduqué.

Diane Halpern et l'approche cognitive à l'EC

Halpern considère l'EC comme l'ensemble des capacités qui sont nécessaires afin de réussir dans des tâches complexes (Halpern 2013[31]). Ces tâches sont assez variées et vont de la prise de décision avec enjeux sociétaux à la formation d'opinions et croyances sur la base des meilleures informations disponibles et à la résolution de problèmes. L'ensemble des capacités de l'EC sont à géométrie variable en raison des contextes. Ces capacités sont cependant toujours dirigées de manière volontaire et consciente (Halpern 1999[32]).

"*Critical thinking refers to the use of cognitive skills or strategies that increase the probability of a desirable outcome. Critical thinking is purposeful, reasoned, and goal-directed. It is the kind of thinking involved in solving problems, formulating inferences, calculating likelihoods, and making decisions. Critical thinkers use these skills appropriately, without prompting, and usually with conscious intent, in a variety of settings. That is, they are predisposed to think critically.*" (Halpern 1999)

Il s'agit donc de capacités de pensée d'ordre supérieur – qui, par définition, sont nécessaires chaque fois que nous nous trouvons face à un problème complexe – alors qu'utiliser des capacités d'ordre inférieur amène à donner probablement une réponse simpliste ou erronée au problème. La spécificité de l'EC réside dans son aspect évaluatif.
La notion d'évaluation est large et comprend l'évaluation de ses propres raisons pour agir d'une certaine manière, l'évaluation des résultats d'un certain processus de pensée, l'évaluation des résultats d'une action dans la résolution de problèmes, l'évaluation de ses propres décisions (Halpern 2013 [33]). Halpern oppose donc EC (pensée, réflexion, apprentissage de concepts) et non-EC (absence de pensée, de réflexion, apprentissage par cœur), mais elle distingue aussi dans l'EC d'autres formes de pensée ou capacités de haut niveau (Halpern 2007[34], 1998[35]).

Halpern fournit plusieurs taxonomies des capacités propres à l'EC. Celles-ci incluent les compétences nécessaires, ou du moins utiles, pour augmenter les chances de résoudre un problème complexe. Les capacités sont divisées en grands groupes (macro-capacités) : compréhension du problème et de son contexte, questionnement ; utilisation des connaissances stockées en mémoire ; compétences relatives au langage, d'analyse du langage utilisé pour porter des arguments ; capacités de raisonnement inductif et déductif, logique en général ; analyse de la structure des arguments ; compréhension des probabilités ; stratégies et outils pour la résolution de problèmes ; pensée créative (Halpern 1999). La taxonomie plus synthétique inclut uniquement cinq grands groupes de capacités : raisonnement verbal, analyse d'arguments, formulation et test d'hypothèses, raisonnement sur la probabilité, décision et résolution de problèmes (Halpern 1998[36]).

L'approche de Halpern présente aussi une dimension normative. Il ne suffit pas en effet d'exercer les capacités citées pour faire preuve d'EC. Certains critères doivent être appris et appliqués, par exemple ceux qui permettent de distinguer les arguments qui sont bien confortés par des preuves (voir Halpern 2013[37]). D'un point de vue pratique, Halpern est optimiste quant à la possibilité d'enseigner l'EC ; elle s'appuie pour cela sur les résultats d'expérimentations à large et à petite échelle (Halpern 1999[38]). Elle est cependant consciente de la difficulté de transférer des capacités à des domaines éloignés, car les capacités acquises peuvent ne pas être activées par le nouveau contexte. Le transfert demande d'avoir accès à la structure profonde du problème, au-delà de son apparence superficielle (Halpern 1999[39]). D'un point de vue opérationnel, Halpern prône donc pour une éducation explicite des capacités de l'EC. Cet enseignement ne serait toutefois pas suffisant et devrait s'accompagner d'un enseignement des dispositions à penser et à apprendre avec effort ; l'enseignement devrait être organisé autour du transfert des capacités apprises et avoir recours à la métacognition explicite.

Halpern propose ainsi une méthode d'enseignement comportant quatre composantes (Halpern 2013, Halpern 1999). (1) La première consiste à apprendre à utiliser les capacités qui permettent de mieux penser et à reconnaître les situations dans lesquelles ces capacités sont requises (Halpern 2013[40]). Mais l'apprentissage ne se limite pas à développer ces capacités ou à en prendre connaissance (par exemple, en lisant un livre qui en parle). (2) Il faut en plus être motivé à les utiliser de manière consciente, et avoir la volonté de contrôler la vérité des affirmations, de chercher des informations pertinentes, de persévérer quand la solution à un problème n'est pas immédiate. Des attitudes ou dispositions que Halpern indique comme des soubassements pour l'EC sont : la volonté de planifier ; la flexibilité au sens de l'ouverture d'esprit qui permet de prendre en considération plusieurs idées différentes et nouvelles ; la

persévérance ; la disponibilité à s'autocorriger, à changer d'idée, à admettre ses erreurs ; l'attitude qui consiste à réfléchir sur soi-même, sur ses propres procédures, plutôt que d'agir en pilote automatique ; la volonté de rechercher un consensus au sein d'un groupe (Halpern 1999[41]). (3) L'apprentissage de l'EC demande également que l'élève soit capable de et disposé à monitorer ses processus de pensée, à en vérifier l'efficacité par rapport aux objectifs, à décider quand allouer des ressources supplémentaires. La métacognition n'est pas pour Halpern une composante de l'EC, mais une fonction d'accompagnement qui facilite le fait de se servir des capacités de l'EC. L'instructeur peut aider l'élève à utiliser sa métacognition en lui posant des questions (Halpern 1999[42]). (4) L'enseignant doit enfin aider l'élève à transférer les compétences générales de l'EC à différents contextes et contenus. Pour cela, il doit multiplier les exemples d'application et expliciter les capacités de pensée requises pour l'EC (Halpern 1999[43]). De manière synthétique, il s'agit de : (1) apprendre explicitement les capacités de l'EC ; (2) développer des dispositions pour une pensée qui demande effort et apprentissage ; (3) apprendre de manière à faciliter le transfert entre domaines ; (4) rendre le monitoring métacognitif explicite et conscient (Halpern 2013[44]).

Deanna Kuhn et l'approche développementale à l'EC

L'approche de Halpern est cognitive – basée sur un ensemble de fonctions ou capacités cognitives – mais elle ne tient pas compte des facteurs du développement. Ainsi, en 1999, Kuhn remarque que les approches éducatives à l'esprit critique – à l'exception peut-être de celle de John Dewey à laquelle elle s'inspire – ne prennent pas suffisamment en compte les connaissances empiriques concernant les trajectoires de développement des facultés cognitives de l'enfant et de l'adolescent (Kuhn 1999[45]).
Comme Dewey, Kuhn considère que le but de l'éducation est de prendre connaissance des capacités et tendances naturelles de l'enfant afin de les développer en attitudes et capacités plus avancées. Pour cela, il est nécessaire de connaître le fonctionnement intellectuel de l'enfant, grâce à des études empiriques. Dans le cadre de ses études sur le développement cognitif, Kuhn s'est plus particulièrement intéressée à la manière dont les enfants réussissent à articuler théorie – leurs connaissances préalables – et données – les nouvelles informations, les preuves – dans leur raisonnement. L'articulation entre faits et théories constitue un aspect fondamental de la science professionnelle, et de la construction de connaissances plus généralement. Kuhn considère cette articulation comme un aspect fondamental de l'EC, défini comme la capacité de distinguer entre connaissances à proprement parler et opinions, de décider quoi croire et de justifier son choix de manière argumentée. Or, selon Kuhn ce type d'articulation exige la mise en place de processus de second ordre, métacognitifs, qui deviennent ainsi la base cognitive développementale de l'EC (Kuhn 1999[46]). Kuhn identifie plus particulièrement trois types de processus métacognitifs impliqués dans l'EC.

Premièrement, la méta-connaissance, c'est-à-dire la réflexion sur la nature des croyances et de la connaissance. Le développement de la capacité à reconnaître l'existence de fausses croyances est ici un élément crucial, car il manifeste la compréhension qu'une croyance n'est pas une simple copie de la réalité. Ce développement permet également de comprendre qu'une assertion est différente des preuves qui peuvent la conforter. À son tour, la compréhension de la notion de preuve permet de se demander : « comment sais-je ce que je sais ? Quelle est la source de ma connaissance ? » Il s'avère que ce type de capacités suit une trajectoire développementale. Par exemple, la compréhension de la distinction entre faits et croyances ou faits et pensées est encore très limitée chez l'enfant de trois ans. Aux alentours

de quatre ans, la capacité à attribuer à autrui des fausses croyances représente une pierre miliaire dans ce développement, car l'enfant est capable – ce que montrent de nombreuses expériences de laboratoire – de reconnaître que ce qu'il pense peut ne pas correspondre à réalité. (Il faut souligner que, selon la littérature récente, le développement des capacités métacognitives, notamment liées aux fausses croyances, est probablement moins linéaire et plus précoce, du moins sous une forme implicite. Voir Beran, Perner, Proust 2012). La capacité à lier positivement assertions et preuves se développe plus tardivement, et avec plus de limites, tout en montrant une variabilité individuelle. Par exemple, les enfants d'âge préscolaire sont souvent encore incapables d'indiquer, quand ils affirment quelque chose, s'ils se basent sur leur propre expérience ou sur le récit de quelqu'un d'autre, ou de préciser à quel moment ils ont acquis cette connaissance et dans quelles circonstances. Si on demande « comment le sais-tu ? », ils ne semblent pas faire la différence entre un constat factuel et une explication basée sur la plausibilité. Par exemple, si on montre à des enfants d'âge préscolaire des vignettes d'une course, présentant un enfant qui porte des chaussures particulièrement « performantes » et, dans l'une des vignettes, cet enfant soulevant un trophée, et qu'on leur demande « comment savez-vous que l'enfant a gagné ? », on constate que, souvent, les réponses ne font pas de distinction entre le fait de voir l'enfant soulever un trophée et l'explication « l'enfant a des chaussures qui font courir vite ». La capacité à réaliser cette distinction se développe avec l'âge, à partir de six ans. Le développement des capacités métacognitives des enfants et adolescents se caractérise par une capacité accrue à coordonner de façon consciente et volontaire théorie et données empiriques, idées préalables et faits ou preuves, aussi bien que par le développement de capacités permettant de monitorer les stratégies utilisées pour gagner de nouvelles connaissances (par exemple, les stratégies plus efficaces pour mémoriser). Kuhn considère que le développement de ces capacités est fondamental pour l'acquisition progressive d'une pensée scientifique. Un point critique est toutefois représenté par le constat que même un adulte n'obtient pas spontanément une réelle maîtrise sur ces capacités. En d'autres mots, ces capacités ne se développent qu'imparfaitement – du moins par rapport à un « idéal » représenté sous la forme de pensée scientifique déployée par les scientifiques de profession. Au cours du développement, en outre, les stratégies moins performantes ne sont pas substituées par celles plus efficaces et sophistiquées. Ce qui change concerne plutôt la fréquence avec laquelle on s'appuie sur l'une plutôt que sur l'autre, sans que le stratégies plus primitives ne disparaissent complètement.
Certaines limites remarquées chez les enfants d'âge préscolaire persistent ainsi chez l'adolescent et l'adulte (Kuhn 1999[47]).

La deuxième composante métacognitive de l'EC identifiée par Kuhn est la métacognition stratégique, qui consiste à se monitorer et à adopter des stratégies plus efficaces pour apprendre et mémoriser.
Enfin, la troisième composante est celle, épistémologique, permettant de comprendre d'où vient la connaissance. Kuhn a une vision « constructiviste » de la connaissance, selon laquelle le réalisme (la connaissance vient de la réalité extérieure) est le niveau le plus bas, d'un point de vue épistémologique. Le réalisme ignorerait les complexités de la construction des connaissances, sa nature en partie subjective, et donc rendrait difficile de comprendre le désaccord entre idées et représentations du monde. On trouve ici selon Kuhn la composante la plus importante, ou de moins la plus spécifique de l'EC. Il s'agit de comprendre que la connaissance n'est ni une copie fidèle de la réalité, ni juste une opinion subjective, mais une articulation entre idées et faits, théories et données. Cette troisième composante comporte une vision de la connaissance et de la science qui n'est pas dépourvue d'une composante culturelle. Or, Kuhn envisage de la connecter comme les autres à un pattern développemental

(Kuhn 1999[48]). En effet, pendant l'enfance et l'adolescence, la composante épistémique évolue comme les autres, mais sans nécessairement atteindre un état épistémologique mûr : celui qui permet de reconnaître à la fois l'objectivité et l'influence des idées sur les connaissances.

Éduquer l'EC permet de maximiser la réalisation du potentiel intellectuel. Il s'agit donc de peaufiner un processus dont la réalisation finale est déjà préconfigurée dans ses débuts. L'éducation de l'EC passe par l'identification claire des composantes développementales de l'EC. Cette identification est d'autant plus importante si les éducateurs doivent insérer leur propre éducation de l'EC dans un contexte riche, disciplinaire, dense de contenu, car il est alors encore plus facile de perdre de vue l'objectif pédagogique. La perspective métacognitive permet également de traiter le problème des dispositions : être disposé à utiliser ses capacités nécessite plus que la simple habitude de le faire, il s'agit de trouver une motivation intrinsèque, et réfléchir sur ses capacités et limites peut permettre cette motivation.
Plus récemment, Kuhn a développé une théorie argumentative de l'EC. Dans ce deuxième volet, Kuhn embrasse une définition assez générale de l'EC, en tant que « forme de raisonnement qui inclut réflexion, justification, application du raisonnement » (Kuhn 2018[49]). Les capacités argumentatives – analyser des arguments, des contre-arguments et comprendre la valeur de différents points de vue – sont alors considérées comme des prérequis pour l'EC ainsi défini.

Kuhn a mené plusieurs études empiriques pour vérifier l'impact de l'argumentation sur l'EC. Cette mesure passe par un entretien à questions ouvertes visant à évaluer la capacité de juger de la crédibilité d'une source, la structure d'un argument et sa qualité, le fait de prendre une position et de prendre en compte des arguments contraires, la conscience métacognitive (Kuhn 2018[50]). Kuhn identifie donc ces capacités avec des capacités de second ordre ou métacognitives : la capacité à penser ce qui constitue une connaissance, la capacité à penser à la manière d'obtenir ou de justifier une connaissance. De façon plus générale, la capacité à s'interroger et à réfléchir sur ce que cela veut dire de « connaître ».
Les interventions proposées par Kuhn sont assez riches, comprenant des composantes métacognitives, du travail de groupe, de la discussion, et de l'instruction explicite sur la manière d'analyser des arguments. Même dans un contexte expérimental impeccable (ce qui est loin d'être le cas pour l'étude présentée, raison pour laquelle les résultats ne sont pas pris en compte ici), il serait donc impossible d'établir si les aspects argumentatifs sont les plus efficaces de la méthode, et si, parmi ces aspects, les plus efficaces sont ceux relatifs à l'instruction directe sur la manière d'analyser des arguments ou plutôt la participation à des groupes de discussion.

Richard Nisbett et le *mindware* – l'outillage pour l'EC

Psychologue social, connu pour ses études sur l'erreur fondamental d'attribution et les influences culturelles sur la pensée, Richard Nisbett ne propose pas de définition de l'EC. Cependant, il s'intéresse à l'enseignement et à l'amélioration de certaines capacités de raisonnement qui sont souvent considérées comme relevant de l'EC, d'où sa place dans ce rapport. Il a en outre mis en place et testé avec succès des enseignements de « capacités de raisonnement ». De plus, il est cité par Diane Halpern comme exemple de la possibilité d'éduquer les « capacités générales », donc l'EC.

Nisbett défend en effet une vision à contre-courant des capacités dites générales : à l'opposé d'une longue tradition remontant à Thorndike, il pense que des capacités générales existent et peuvent être éduquées avec succès, le succès consistant notamment dans la possibilité de transférer les acquis à des contextes et contenus éloignés, pourvu que les règles de résolution soient les mêmes. Ainsi, Nisbett et ses collègues ont testé les effets de l'instruction d'un groupe de sujets à la loi des grands nombres. Des études préalables avaient déjà montré que cette règle était généralement comprise – comme règle abstraite, au sens qu'elle s'appliquait à une variété de contextes – par une personne lambda. Cependant, elle n'était pas adoptée dans certains contextes, notamment ceux qui se prêtaient moins à être considérés comme quantitatifs. L'instruction explicite à la règle et l'instruction implicite *via* des exemples d'application ont toutes les deux donné des résultats positifs sur l'application spontanée et le transfert à des contextes différents de ceux de départ. Ceci confirmerait, selon Nisbett, l'existence d'une règle générale (Nisbett et al. 1987[51]). Mais l'optimisme de Nisbett quant à l'éducabilité de quelque chose comme « la raison » en général ou l'EC est fortement limité par le constat selon lequel les capacités générales qui se comportent ainsi ne sont pas nombreuses. Par exemple, les règles de la déduction logique ne se comportent pas comme les règles statistiques (Nisbett et al. 1987[52]). Nisbett en tire la conclusion que seules les règles abstraites comme celles statistiques peuvent être entraînées de façon générale. Il s'agit donc d'identifier les autres règles générales qui se comportent de la sorte. Une fois réunies, celles-ci constituent un programme d'éducation formelle avec des applications dans toute une variété de contextes.

Dans son livre dédié à outiller l'esprit humain pour le rendre plus rationnel et efficace, Nisbett propose une boîte à outils cognitive (*mindware*) composée d'environ une centaine de concepts, principes, règles d'inférence développés par des psychologues, statisticiens, philosophes, logiciens. Il s'agit d'outils qui s'ajoutent au sens commun, aux intuitions naturelles (*supplements to common-sense*) et qui sont censés permettre de penser et d'agir de façon plus efficace. Notre sens commun nous amène en effet, parfois, à commettre des erreurs de jugement. Nisbett souligne ainsi que la plus grande partie de nos processus cognitifs sont des inférences silencieuses basées sur des connaissances tacites. Ces inférences sont le fruit de l'évolution qui nous a dotés de mécanismes pour catégoriser, expliquer, argumenter, etc. L'exemple classique est celui des illusions, par exemple les illusions de perspective : notre évolution s'est faite dans un monde tridimensionnel et inférer certaines propriétés des objets à partir de leur position dans l'espace facilite la prise de décision. Des dessins bidimensionnels peuvent cependant nous induire en erreur. Ces considérations s'appliquent aussi au raisonnement sur les catégories (reconnaissance et catégorisation d'objets et d'individus), sur les causes, etc. Le résultat est une divergence entre nos « perceptions » de la réalité et la réalité elle-même (Nisbett 2015[53]). Nisbett évoque notamment les biais classiques décrits par des psychologues tels que Daniel Kahnemann et Amos Tversky comme la représentativité, le *framing*, l'ignorance du hasard et la perception surdimensionnée de patterns, la disponibilité. De façon plus générale, Nisbett liste les causes d'erreur de jugement suivantes :

- nous ignorons que des facteurs sans pertinence avec le problème influencent silencieusement nos jugements (voir les biais cités) ;
- nous sommes ignorants du contexte et considérons peu son rôle dans le comportement des individus (exemples : erreur fondamental d'attribution, effets de l'influence sociale) ;

- nous ignorons que nos choix sont influencés par des éléments dont nous n'avons pas une perception consciente, nous ignorons les causes réelles de nos opinions, choix, décisions ;
- nous ne sommes pas conscients de l'importance de mener une analyse coûts-bénéfices avant de prendre une décision, de l'influence de biais comme l'aversion à la perte ;
- nous ignorons que beaucoup de situations de la vie quotidienne peuvent être décrites selon des règles statistiques de fréquence, dispersion, moyenne et régression à la moyenne ;
- nous commettons des erreurs de corrélation et causalité illusoires, encore une fois liées à notre ignorance des règles statistiques d'association ;
- nous faisons confiance à nos suppositions, et donc nous ne les testons pas de façon scientifique (test A/B avec randomisation ou sans, avec toutes les considérations statistiques allant avec la significativité, covariance, etc.).

Les principes et règles proposés par Nisbett sont censés agir comme correctifs. Selon lui, ils peuvent être appris, devenir automatiques et être utilisés sans effort dans une variété de problèmes de la vie quotidienne (Nisbett 2015[54]). La solution proposée par Nisbett consiste essentiellement à prendre conscience du fonctionnement de la cognition humaine et de ces biais et heuristiques. Cette prise de conscience devrait déjà nous rendre plus humbles (baisser notre confiance en nous-mêmes et en nos jugements) (Nisbett 2015[55]). Grâce à cela, il devient possible d'avoir recours à des outils cognitifs plus adaptés à la situation. Pour pouvoir appliquer les bons outils (règles, principes) au bon moment, un effort ultérieur doit être réalisé pour :

- formuler et encadrer (*frame*) le problème afin que l'utilisation de la règle devienne immédiatement pertinente ;
- codifier le problème afin que les principes puissent y être appliqués. Dans le cas des règles statistiques, l'idée de Nisbett semble être de familiariser le plus grand nombre avec une manière de penser statistique appliquée à la vie de tous les jours (Nisbett 2015[56]).

L'approche proposée par Nisbett peut soulever plusieurs critiques, notamment car les enseignements proposés semblent parfois difficiles à réellement importer dans la vie quotidienne.

Tout d'abord, la nécessité de coder le problème à son « bon niveau » d'abstraction et de l'encadrer afin de rendre saillant le type de principe capable de le résoudre renvoie directement à la littérature sur le transfert et sur les difficultés du transfert que Nisbett critique. La suggestion ne permet pas de savoir, par exemple, à quels éléments du contexte nous devons faire attention, quels éléments nous influencent silencieusement dans un contexte donné...

Ensuite, pourquoi, si ces principes sont tellement universels, nous trompons-nous si systématiquement ? Qu'est-ce qui rend notre fonctionnement cognitif, pourtant capable d'adopter des règles générales, au fond réfractaire à le faire – du moins en l'absence d'une éducation dédiée ? Ainsi, lorsque Nisbett suggère que nous devrions percevoir les questions relatives à la personnalité et au comportement comme des problèmes d'échantillonnage, il néglige d'expliquer pourquoi nous ne le faisons pas spontanément. Que faisons-nous en réalité et pourquoi ? En pratique, l'approche de Nisbett ne semble pas suffisante pour expliquer le décalage entre notre fonctionnement naturel et un fonctionnement augmenté qui semble tout de même bien plus adaptatif que le premier.
En troisième lieu, l'idée de prendre conscience de notre fonctionnement et de nos biais pour devenir plus humble n'est ni réaliste ni nécessairement souhaitable. Nisbett approuvent l'idée des penseurs de l'EC qui semblent considérer positif que nous doutions systématiquement de

nous-mêmes, alors que cette attitude aurait des conséquences désastreuses sur notre capacité d'agir. De plus, elle est justement assez contraire à l'idée selon laquelle l'évolution nous a dotés de solutions prêtes à l'emploi souvent utiles et efficaces. En réalité – sans nier que notre fonctionnement, nos heuristiques et nos biais nous induisent parfois en erreur –, ce que nous aimerions souvent connaître concerne le moment où un phénomène a des chances d'arriver, afin d'apprendre à identifier et à anticiper les situations à risque.

La notion d'outils nous rendant plus intelligents est une notion cependant très intéressante. D'après Nisbett, nous sommes devenus plus intelligents au cours de notre évolution culturelle. Nous avons inventé des outils mathématiques et statistiques, la méthode scientifique, la notion d'hypothèse, de test. Ces outils sont enseignés systématiquement mais, normalement, ils ne le sont pas en lien avec l'idée de développement de l'EC et ne sont pas enseignés de manière à faciliter leur utilisation dans la vie quotidienne.

Keith Stanovich : EC = rationalité

Keith Stanovich est l'un des représentants majeurs du courant des modèles cognitifs « duals » (*dual system models*). Dans ces modèles est postulée une organisation cognitive en deux grandes typologies de processus, parfois appelés « systèmes » : les processus de type 1 et les processus de type 2, ou Système 1 (S1) et Système 2 (S2). S1 est une modalité de fonctionnement (non un système au sens d'une architecture cognitive particulière, d'un module) basé sur des heuristiques générales et rapides de résolution de problèmes ou de jugement ; S2 est une modalité de fonctionnement qui se base sur l'implémentation d'algorithmes *ad hoc* pour la situation, choisis de façon délibérée. Selon Stanovich, les deux systèmes opèrent en parallèle : le Système 1 est toujours actif, et le Système 2 intervient ou prend le dessus sur le Système 1 dans certaines occasions. En réalité, il existe une variété de modèles dualistes, et chaque modèle a subi des changements au cours du temps (Evans & Stanovich 2013[57]). Stanovich a même ajouté un troisième système, S3, qui naît de la décomposition du S2 en un système analytique et en un système réflexif. S3 entre en jeu quand un conflit entre S1 et S2 algorithmique est détecté ; son but est de résoudre le conflit (Evans & Stanovich 2013[58]).

La théorie de l'EC proposée par Stanovich se situe dans ce cadre. Pour lui, l'étude de l'EC a nécessairement une dimension normative : son objectif est d'enseigner et d'apprendre à mieux penser (Stanovich & Stanovich 2010[59]). Mais que signifie « mieux penser » ? Et pourquoi mieux penser serait positif ? Stanovich résout la question en reliant l'idée de mieux penser avec celle de rationalité, de pensée rationnelle. La pensée rationnelle est celle qui dispose du plus de chances de nous amener à former des croyances en adéquation avec la réalité et nous permettant d'atteindre nos objectifs. Des capacités ou attitudes comme l'ouverture d'esprit sont au service de cette rationalité, mais ne constituent pas son cœur. Stanovich établit ainsi une hiérarchie entre les capacités de l'EC, avec un objectif final à atteindre qui est celui de la rationalité, et des capacités qui peuvent constituer des moyens ou des capacités associées. Si nous estimons l'ouverture d'esprit et d'autres capacités indiquées dans la littérature comme propres à l'EC, alors nous devons considérer l'éducation de l'EC comme ayant un but plus « élevé » : le développement d'une pensée rationnelle. De façon plus générale, il s'agit de lier le concept d'EC à celui de rationalité. Cette association entre rationalité et EC comporte, selon Stanovich, plusieurs avantages, dont celui d'ancrer le concept relativement flou d'EC sur une notion plus développée dans le cadre de la recherche en sciences cognitives, tant au niveau théorique que des données empiriques (Stanovich & Stanovich 2010[60]).

Dans sa définition de la rationalité (et donc de l'EC), Stanovich cite deux composantes, une épistémique et une pragmatique ou instrumentale. Les deux sont étudiées de manière étendue en sciences cognitives. D'un côté, la rationalité épistémique est définie en relation avec les concepts de confiance bien calibrée et de preuves disponibles : le degré de confiance dans une affirmation est bien calibré par rapport aux preuves disponibles et pertinents. De l'autre, la

rationalité pragmatique est associée à la capacité de satisfaire de manière optimale ses objectifs. Les deux formes de rationalité sont liées, la première étant au service de la deuxième. La rationalité nous permettrait donc e maximiser l'utilité, grâce à la possibilité de suivre certains patterns de choix rationnels (Stanovich & Stanovich 2010[61]). Or, selon la théorie des heuristiques et biais, la rationalité est souvent enfreinte au quotidien, à cause de nombreux biais documentés dans la littérature. Pour Stanovich, donc, les biais sont des influences néfastes qui amènent à s'éloigner d'une rationalité « idéale » et idéalement souhaitable à l'optimisation des choix. Stanovich ne prend ici pas en considération ni les raisons évolutives des biais ni le fait que les heuristiques dont il est question ont le plus souvent un impact positif sur la décision, qu'elles rendent optimale dans un contexte précis (Stanovich & Stanovich 2010[62]). Deux notions importantes dans le cadre de la théorie de l'EC/rationalité proposée par Stanovich sont donc celles de croyances préalables et de biais. Pour Stanovich, la caractéristique essentielle de l'EC est la capacité à évaluer preuves et arguments indépendamment de nos propres croyances préalables, des opinions, de tous les biais ou associations qui nous amènent à donner des réponses automatiques, non conscientes, non réflexives, en relation avec le contexte. En d'autres termes, il s'agit d'une pensée non biaisée, capable de de-contextualisation. Stanovich souligne que cette capacité est présente dans plusieurs approches de l'EC, tout comme dans différentes théories de la pensée.

Ce qui manque à ces approches et qui doit être intégré à une théorie de l'EC en tant que rationalité concerne la prise en compte de la littérature sur les heuristiques et les biais, en particulier les tests pouvant servir de tests de l'EC/rationalité (Stanovich & Stanovich 2010[63] ; West et al. 2008[64]). L'approche proposée par Stanovich permet ainsi d'envisager des idées de mesure de l'EC/rationalité : il s'agit d'évaluer le degré de résistance aux biais, d'indépendance de la pensée par rapport à ces influences et à l'influence des croyances préalables, mais aussi d'évaluer la résistance à d'autres formes de biais décrits dans la littérature sur les heuristiques et les biais (Stanovich & Stanovich 2010[65], West, Topiak & Stanovich 2008[66]). Développer l'EC – c'est-à-dire développer la rationalité – signifie donc développer et renforcer les processus de pensée de type 2, et évaluer l'EC signifie donc évaluer la mesure dans laquelle le sujet a plutôt recours aux processus de type 2 et est capable de mettre sous contrôle les processus de type 1. En effet, selon Stanovich, l'approche des heuristiques et biais, avec tout son appui empirique, dépend d'un point de vue théorique d'une approche duale de la pensée. Dans le cadre de cette dernière, les processus de type 2 – coûteux d'un point de vue cognitif, sériels, faisant l'objet d'accès conscient, souvent verbalisés – ont comme fonction de dépasser les processus de type 1 – rapides, parallèles, peu coûteux, basés sur des associations anciennes et automatiques, et pouvant produire des réponses irrationnelles en raison du contexte dans lequel ils sont activés.

Une fonction cognitive fondamentale pour l'exercice de la rationalité devient donc celle d'« inhibition », étudiée dans le cadre des études concernant les fonctions exécutives. L'inhibition est le mécanisme qui permet aux processus de type 2 de mettre sous contrôle ceux de type 1 et de prendre leur place (Stanovich & Stanovich 2010[67]).

De sa théorie duale, Stanovich tire une considération ultérieure concernant la rationalité/EC. Rationalité et EC appartiendraient à un domaine différent de l'intelligence telle que les tests standardisés actuels la mesurent. En effet, les mesures de l'intelligence tendent à se baser sur des mesures de l'efficience dans l'élaboration de l'information (*information processi ng*) dans des situations optimales : le sujet sait qu'il doit donner la meilleure réponse possible et donc utiliser ses meilleurs algorithmes. Les tests d'intelligence le mettent donc implicitement en situation d'utiliser ses processus de type 2, et mesurent ensuite ses performances, son efficacité dans cette mise en place. En revanche, un test de rationalité est un test conduit dans une situation typique (vs optimale) qui permet d'évaluer si le sujet met en place des processus de type 2 alors que la situation permet, voire oriente vers, l'utilisation des processus automatiques et peu coûteux. Les tests de rationalité doivent mesurer cette volonté d'effort cognitif, et la disposition à penser de manière rationnelle : par exemple, si le sujet a tendance à chercher des informations avant de se forger une opinion, à calibrer la force de ses opinions

sur la base de celle des preuves disponibles, à songer aux conséquences, à réfléchir avant de donner une réponse, à peser le pour et le contre d'une décision, à s'autoréguler par rapport à ses objectifs. À l'inverse, les tests d'intelligence ne concernent pas la capacité de changer d'idée face à des preuves contraires, à diriger les actions de manière à atteindre un objectif, ou à obtenir de nouvelles connaissances (Stanovich & Stanovich 2010[68]).

Stanovich a ainsi reparti les processus de type 2 en deux typologies ultérieures : pensée de type algorithmique et pensée réflexive. La première est caractérisée par la capacité à mettre en place des processus de type algorithmique, la seconde par la disposition à le faire, c'est-à-dire à prendre une attitude réflexive face à un nouveau problème. Selon Stanovich, plusieurs études montrent que les résultats à des tests cognitifs de type QI et à des tests de résistance aux biais, notamment au *belief biais* (le biais des croyances préalables), sont décorrélés. Donc que nous ne pouvons pas attribuer le succès ou l'insuccès dans les tests de rationalité uniquement à la mesure de l'intelligence (Stanovich & Stanovich 2010[69]).
L'exercice de la rationalité requiert pour Stanovich trois formes de capacités différentes :

- les capacités algorithmiques doivent être présentes ;
- une tendance à bloquer et dépasser les réponses automatiques doit être présente quand celles-ci sont suboptimales ;
- il faut posséder des connaissances qui peuvent être récupérées en mémoire et qui vont aider la résolution de problèmes (ce que Stanovich appelle « *mindware* »).

Or, les tests de QI se contentent de mesurer la première forme de ces capacités. De bons résultats dans ce type de test n'excluent pas des résultats pauvres en termes de rationalité, par exemple lors de tests à base de biais comme le « test de la batte et de la balle ». Stanovich parle alors de « dysrationalité » : l'incapacité d'être rationnel malgré un QI normal. Les raisons sont des défauts concernant les autres capacités nécessaires à la rationalité : le sujet préfère des solutions moins efficaces mais aussi moins coûteuses, ou il ne possède pas le *mindware* nécessaire (Stanovich & Stanovich 2010[70]).

Plusieurs aspects de l'approche proposée par Stanovich sont originaux par rapport aux autres théories de l'EC. En particulier, la notion de rationalité épistémique est convaincante pour une définition restreinte de l'EC. L'idée de manipulation de l'environnement pour suppléer aux déficits de motivation est également originale et ouvre la voie à une vision élargie de la rationalité – distribuée sur le sujet et son environnement. Cependant, la théorie duale de la pensée et le programme des heuristiques et biais qui sont à la base de l'approche de Stanovich sont sujets à plusieurs critiques : la rationalité décrite par Stanovich serait une rationalité idéale, faisant abstraction des raisons qui font du fonctionnement cognitif normal un fonctionnement globalement bien adapté pour résoudre les problèmes de la vie quotidienne. La difficulté à trouver la motivation pour dépasser les « erreurs » du système 1 et à leur substituer – *via* la réflexion – les processus plus algorithmiques de type 2 pourrait d'ailleurs être expliquée par cette fonctionnalité globale des solutions rapidement disponibles. Stanovich ne considère pas qu'il soit possible que des systèmes et mécanismes d'autocorrection des erreurs soient aussi présents de façon automatique, implicite, et qu'il pourrait être utile de les développer (sans nécessairement passer par des procédures explicites, réflexives). Stanovich met l'accent sur l'importance des connaissances et des stratégies pour l'exercice de la rationalité/EC. Il semble cependant considérer ces connaissances et stratégies comme générales, applicables dans une variété de contextes et non dépendantes du contenu de domaine. La difficulté à transférer une stratégie d'un contenu ou contexte à un autre n'est donc pas prise en considération par Stanovich. De plus, elle pourrait à la limite être considérée – dans le cadre de son approche – comme un biais à dépasser, une influence du contexte qui empêcherait de généraliser de manière appropriée, rationnelle. Si nous songeons à des stratégies pour prendre en compte la position opposée, ou prendre en compte le principe de falsification, nous pouvons constater que leur application peut être facilitée ou au contraire

rendue très ardue par le fait de posséder ou ne pas posséder des connaissances de domaine approfondies.

Robert Sternberg : EC = intelligence

Sternberg attribue la naissance du mouvement moderne de la pensée critique à John Dewey et célèbre la richesse de ce mouvement, qui se nourrit de trois traditions différentes : celle philosophique, celle psychologique et celle éducative. La tradition philosophique de l'EC puiserait ses racines dans la Grèce antique, plus récemment dans le travail de Ennis, Lipman, Paul... Cependant, Sternberg décrit les approches philosophiques comme étant plus centrées sur les lois normatives des systèmes logiques formels que sur des capacités réelles. Dans ce sens, il considère les approches philosophiques utiles comme guides mais également comme limitées d'un point de vue pratique. La tradition psychologique comprend – pour Sternberg lui-même et d'autres psychologues fortement engagés dans l'éducation, comme Bransford, Feuerstein et Bruner, qui pourtant n'ont pas directement traité la question de l'EC – plutôt celle des capacités d'ordre supérieur ou des formes de pensée. L'utilité de l'approche psychologique réside pour Sternberg dans sa capacité à éclairer les processus mentaux que les humains utilisent quand ils cherchent à penser de manière critique même s'ils sont dans des conditions suboptimales : l'information est limitée, le temps est limité, leur mémoire n'est pas parfaite…

Sternberg apprécie cependant les apports de l'approche éducative (qu'il associe à des figures telles que Bloom) qui ont été capables de proposer des théories de la pensée et de l'apprentissage prenant en compte des observations écologiques, en classe, et pas uniquement (comme c'est le cas pour les psychologues) les résultats de tests quelque peu artificiels conduits en laboratoire.

Sternberg propose une taxonomie des capacités de l'EC basée sur une approche psychologique, en lien avec sa théorie de l'intelligence. Cette taxonomie comprend trois types de capacités impliquées dans l'EC : des méta-composantes, des performances, des processus d'acquisition de connaissances. Les méta-composantes sont des processus d'ordre supérieur utilisés pour planifier, monitorer ce qui est fait et évaluer les résultats de ce qui a été fait. Par exemple, ces composantes comprennent la capacité d'identifier l'existence d'un problème, de comprendre sa nature, de planifier de manière ordonnée dans le but de le résoudre, de créer une stratégie cohérente, etc. (Sternberg 1986[71]). À ces composantes d'ordre supérieur s'ajoutent différentes formes de raisonnement et des capacités qui permettent de mettre en pratique les instructions des composantes métacognitives : raisonnement spatial, causal, déductif, capacité de lecture… Les composantes de performance sont des processus d'ordre inférieur. Ces capacités interviennent de manière variable selon la situation. Une liste exhaustive ne pourrait pas être produite. Par exemple, il existe des composantes spécifiques qui permettent de lire, d'autres qui permettent de raisonner par induction, etc. (Sternberg 1986[72]). Le troisième pilier cognitif de l'EC est représenté par les capacités qui permettent d'acquérir de nouvelles connaissances ou d'apprendre de nouvelles procédures. L'acquisition de connaissances et procédures se base à son tour sur trois composantes principales : l'encodage sélectif, la combinaison des informations de manière sélective et pertinente, la comparaison entre connaissances acquises et nouvelle information à apprendre. Ces mêmes composantes sont celles qui permettent de faire face à la nouveauté, une condition fondamentale d'exercice de l'EC (Sternberg 1986[73]).

L'EC est donc un ensemble de processus mentaux, de stratégies, de représentations que nous utilisons pour résoudre des problèmes, prendre des décisions, apprendre des concepts nouveaux. Au-delà des éléments plus généraux, métacognitifs, les capacités nécessaires à l'EC varient en fonction de la tâche, de la situation et aussi des sujets. D'un point de vue éducatif pratique, Sternberg est conscient de l'importance de favoriser le transfert des capacités acquises. Ceci se ferait grâce à une pédagogie qui maximiserait les chances

d'appliquer les apprentissages à des situations de la vie réelle. Au niveau pratique, il propose d'utiliser des exemples concrets pour illustrer les méthodes de la pensée critique, des exercices variés pour l'entraînement, comportant des exercices pratiques sur des questions quotidiennes mais aussi sur des questions plus académiques. Dans un manuel pédagogique, il propose de façon plus générale d'améliorer les capacités intellectuelles des élèves, qu'il appelle « intelligence ». Le manuel s'oriente vers les trois composantes de l'approche décrite mais comprend également des parties concernant les aspects émotionnels et la motivation à utiliser son intelligence. Au niveau des tests pour mesurer l'EC, Sternberg cite le *Watson-Glaser Critical Thinking Appraisal*, le *Cornell Critical Thinking Test*, le *New Jersey Test of Reasoning Skills*, tous les trois dérivés de la tradition philosophique. Il développe son propre test psychologique : *the Triarchic Test of Intellectual Skills*. Ce test ne sépare pas, selon l'approche propre à Sternberg, EC et intelligence et cherche à mesurer les composantes méta (planifier, monitorer, évaluer), les composantes performatives (inférer des relations, appliquer des relations, comparer des relations d'ordre supérieur entre domaines), les composantes d'acquisition des connaissances et la capacité à faire face à la nouveauté (distinguer informations pertinentes et non pertinentes, combiner les informations de manière logique, utiliser les connaissances acquises pour apprendre ou comprendre de nouveaux concepts), l'automatisation des processus d'élaboration de l'information et la flexibilité adaptative.

Tim Van Gelder : l'EC comme forme d'expertise

Van Gelder, philosophe de formation, s'est beaucoup intéressé aux sciences cognitives. Il travaille en tant que consultant pour des entreprises et comme chercheur à l'université de Melbourne. Même s'il n'a publié qu'un seul article sur le thème de l'EC, dans lequel il affirme s'inspirer des sciences cognitives pour en tirer six leçons sur l'enseignement de l'EC, cet article est cependant souvent cité dans les références concernant l'approche psychologique de l'EC. Van Gelder souligne à la fois l'intérêt et les limites de se tourner vers les sciences cognitives pour indiquer comment enseigner l'EC : les sciences cognitives permettent de mieux comprendre le fonctionnement de l'esprit, mais il s'agit d'un domaine de recherche fondamental qui ne fournit pas des indications pratiques et des données empiriques sur ce qui fonctionne ou non ; et les sciences cognitives se sont, selon lui (en 2005), peu occupées des capacités de pensée d'ordre supérieur, et notamment de l'EC.

Les leçons que nous pouvons en tirer sont donc, sauf une, d'ordre assez général : la nécessité de pratiquer les capacités de l'EC mais aussi celle d'acquérir des connaissances théoriques, la difficulté d'acquérir une expertise et la nécessité de pratiquer, pour le transfert, l'existence du biais de confirmation. La seule indication spécifique concerne l'utilité de pratiquer des techniques comme la représentation visuelle des arguments, une pratique que Van Gelder encourage par ailleurs (Van Gelder 2005[74]). La théorie fondamentale de Van Gelder est que l'EC n'est pas naturel. Au contraire, nous pourrions le comparer à la pratique du ballet : une activité fortement artificielle, experte, qui exige un exercice de longue durée, dédié, coûteux. Rien dans notre évolution n'a favorisé le genre de raisonnement rationnel, logique, qui est propre à l'EC (Van Gelder 2005[75]). La position de Van Gelder est certes quelque peu caricaturale, car l'EC pourrait comme d'autres capacités avoir des bases naturelles qui rendraient possible le développement ultérieur, éventuellement culturel de capacités expertes. Le ballet est pour cela un bon exemple car il représente une évolution culturelle de capacités de mouvement et de tendances qui s'expriment naturellement. Le « juste combien » il est utile d'être doté de capacités d'EC pour survivre dans les environnements de notre évolution n'est pas une question à laquelle nous pouvons répondre de manière abstraite, et dépend également des capacités que nous considérons comme fondamentales pour l'EC. Or, Van Gelder ne fournit aucune définition de l'EC.

Une difficulté ultérieure identifiée par Van Gelder dans l'exercice de l'EC est qu'il s'agit d'une capacité d'ordre supérieur (il compare alors l'EC et jouer au tennis), qui se base sur l'exercice d'autres capacités plus basiques. Par exemple, pour exercer son EC dans

l'interprétation d'une lettre, il faut savoir lire, etc. Ces difficultés expliquent pourquoi apprendre à penser de manière critique demande beaucoup de temps, et Van Gelder compare ce au temps nécessaire pour maîtriser une seconde langue. Pour ces différentes raisons, il n'existe pas de potion magique pour devenir de meilleurs penseurs critiques. Cependant, Van Gelder suggère plusieurs stratégies. La première est la pratique. Il s'inspire de la littérature sur l'expertise, notamment l'œuvre de Karl Anders Ericsson, et intègre donc dans le champ de l'EC les conseils d'Ericsson pour développer une expertise dans des domaines aussi différents que le jeu d'échecs, le sport, la pratique d'un instrument musical. Notamment, la pratique doit être délibérée : assumée et volontaire. Elle doit être conçue pour le transfert : par exemple, il s'agit de pratiquer dans des contextes différents des exercices riches en contenus, mais ensuite de favoriser l'abstraction pour identifier les traits communs – même s'il n'y a pas de recette sur la manière de favoriser le transfert. Le seconde est la théorie : par exemple, maîtriser le vocabulaire, les principes de l'argumentation, les lois de la logique. La théorie permet de développer une compréhension de ce qui est fait dans la pratique. Cela signifie essentiellement de dédier des cours à l'EC, qui comprennent théorie et pratique à parts égales. Van Gelder est convaincu des avantages existant à représenter spatialement des arguments afin de mieux en maîtriser les relations, en les visualisant. Il suggère donc d'utiliser ce type d'outil externe pour améliorer les capacités d'EC au niveau des capacités d'argumentation. Enfin, Van Gelder attire l'attention sur une difficulté supplémentaire par rapport à celles évoquées dans son premier point : nous sommes la proie d'illusions, de biais, d'erreurs cognitives qui se produisent de façon silencieuse et appartiennent à notre appareillage cognitif, pour des raisons naturelles mais pouvant être renforcés culturellement. Le biais le plus profond et pervasif est, selon Van Gelder, celui qui amène à préserver ses croyances et opinions, appelé « biais de confirmation ».

Daniel Willingham : l'EC dépend du contenu

Daniel Willingham est psychologue cognitif, expert en éducation. Son analyse de l'EC est souvent citée comme une approche pessimiste quant à la possibilité d'éduquer cette capacité. Willingham est en effet – dans le cadre du panorama que nous venons d'esquisser – le seul théoricien de l'esprit critique fortement sceptique quant à la possibilité d'améliorer ce dernier, ou du moins de pouvoir l'améliorer de manière générale et par la voie d'enseignements généraux. L'approche la plus proche de cette vision parmi celles rencontrées jusqu'ici est représentée par l'œuvre des philosophes Sharon Bailin et Mark Battersby (Bailin & Battersby 2018). Les deux philosophes, comme Willingham, mettent également en effet fortement l'accent sur la dépendance des processus de pensée des contenus sur lesquels la pensée s'engage. Leur approche de l'éducation de l'esprit critique est donc disciplinaire et vise à améliorer les capacités de pensée dans un domaine, ou du moins en relation avec certains contenus. Bailin et Battersby critiquent la volonté d'améliorer l'EC *via* des cours de logique formelle en faveur de la mise en place de débats argumentatifs sur des thèmes spécifiques. Il ne s'agit pas selon eux d'apprendre à se défendre (EC comme autodéfense intellectuelle) contre des arguments mal formulés mais d'apprendre à rechercher les faits derrière les informations, à choisir ses sources, à utiliser les normes épistémiques adaptées au contexte et aux contenus en discussion, à ajuster sa confiance sur la base de ces évaluations. C'est notamment en comparant des positions différentes, *via* le débat et la discussion – une méthode que Bailin et Battersby appellent « *inquiry* », « investigation » – que l'EC peut être amélioré. Les auteurs soulignent que l'investigation est contextuelle : l'analyse des arguments dépend du contenu des arguments, car, par exemple, les normes et les standards pour évaluer la justesse d'un argument peuvent ne pas être les mêmes d'un contenu à un autre. L'EC est donc la capacité à utiliser les normes épistémiques appropriées au contexte et au contenu pour défendre ou attaquer un argument, se justifier ou critiquer une position. Les normes en question permettent d'évaluer la qualité des sources contre des standards, de juger des arguments causaux ou des raisonnements qui impliquent des statistiques (Bailin & Battersby 2018).

Willingham adopte une définition plus descriptive de l'EC, basée sur les capacités cognitives mobilisées. De même, ses raisons pour maintenir une position sceptique face à un enseignement général de l'EC dérivent de considérations d'ordre psychologique, plutôt qu'épistémique.

Pour lui, les capacités de l'EC appartiennent à trois types d'activités mentales que nous utilisons au quotidien, mais pas nécessairement de manière critique : raisonner, résoudre des problèmes, prendre des décisions. L'adjectif « critique » indique que les processus de raisonnement, décision, résolution de problèmes sont menés de manière volontaire (autonome), efficace (non affectée par des biais et autres limites cognitives naturelles) et originale (il ne s'agit pas de se rappeler d'une solution) (Willingham 2007[76]). Le problème est donc : pouvons-nous enseigner à penser ? La réponse est : si nous le pouvions, ce ne serait pas comme l'enseignement ou l'apprentissage du vélo (Willingham 2007[77]). Willingham est pessimiste concernant la possibilité d'enseigner l'EC car il considère que la pensée n'est pas une capacité (*skill*). Au contraire, il explique que la pensée est fortement dépendante des contenus de la connaissance. Par conséquent, apprendre des stratégies générales ne garantit pas de savoir les appliquer dans des cas concrets, sur des contenus particuliers (Willingham 2007[78]). La dépendance de la pensée du contenu se révèle aussi dans la difficulté à se détacher des aspects superficiels, concrets d'un problème (son contenu factuel) pour chercher à entrevoir sa « structure profonde ». Quand nous cherchons à comprendre un nouveau problème, nous utilisons les connaissances stockées dans notre mémoire, mais aussi le contexte. Ceci rend la compréhension plus rapide, mais l'ancre au contenu contextuel. Ce sont ces considérations qui amènent Daniel Willingham à affirmer qu'enseigner – de façon générale – à penser ou à penser de manière critique est impossible.

Pourtant, plusieurs programmes ont pour objectif de développer l'esprit critique ou la pensée de façon générale. Willingham se demande ce qu'il en est de leurs résultats. Même quand ces méthodes semblent fournir des résultats positifs, les évaluations empiriques présentent plusieurs limites de nature méthodologique : si nous pouvons dire de plusieurs de ces méthodes que leurs objectifs « internes » sont atteints (les apprenants apprennent à résoudre le type de problèmes qu'ils rencontrent dans le cadre du programme), il est plus difficile de se prononcer sur la généralisation et le transfert des effets de ces méthodes dans la « vraie vie », ou sur la possibilité qu'ils perdurent dans le temps. Dans certains cas, il est aussi difficile de savoir si l'effet positif est dû à la méthode elle-même ou à d'autres conditions. Willingham propose une rapide analyse de l'efficacité des programmes plus connus d'enseignement de l'esprit critique *per se*. Il synthétise les caractéristiques communes aux méthodes les plus diffusées en trois points :

- elles présument l'existence de capacités (*skills*) pouvant être pratiquées indépendamment du contexte et du contenu, l'enseignement de l'EC a lieu en dehors d'un cadre disciplinaire ;
- certaines ont une durée longue (trois ans, plusieurs heures d'instruction par semaine) ;
- elles utilisent toutes des exemples de pensée critique et ensuite demandent d'appliquer les stratégies apprises ; certaines utilisent des problèmes abstraits (*Ruven Feuerstein Instrumental Enrichment*), d'autres des histoires à mystères (*Martin Covington Productive Thinking*) ou encore des discussion de groupe sur des problèmes quotidiens (*Edward De Bono Cognitive Research Trust* : CORT).

Les études qui mesurent les effets de ces interventions présentent plusieurs limites (Willingham 2007[79]) :

- dans certains cas, les étudiants sont évalués seulement une fois, donc nous ne pouvons pas savoir si les effets perdurent ;
- dans certains cas, il n'existe pas de groupe témoin, ou le groupe témoin ne mène pas une activité alternative (groupe passif) ;

- dans certains cas, il n'y a pas de mesure de transfert à des situations réelles ou à des situations différentes de celles utilisées dans le cadre de l'instruction ;
- seule une petite partie de ces études a été soumise au processus scientifique de publication et de validation par les pairs ;
- quand nous parlons d'effet positifs de ces méthodes, nous ne parlons pas nécessairement de la même chose que d'esprit critique, ou d'une notion qui ait spécifiquement trait à l'esprit critique.

En dépit de toutes ces difficultés et limites, les analyses de la littérature ou les textes sur l'esprit critique expliquent souvent que les interventions éducatives pour l'esprit critique ont des effets positifs (Willingham 2007[80]). N'existe-t-il aucun espoir de transférer des acquis d'un contexte à un autre – y compris des acquis « généraux » comme la capacité à résoudre un certain type de problème ? Willingham indique deux conditions de réussite : la familiarité avec le contenu profond et le fait de savoir qu'il faut regarder le contenu profond. La familiarité dépend d'une pratique répétée, et automatisée. La deuxième stratégie est métacognitive : savoir où et comment chercher dans sa mémoire les bonnes stratégies, ce qui présuppose qu'il est au moins nécessaire de vouloir et de penser à le faire. Il faut enfin posséder des connaissances et savoir les mobiliser dans la pratique. Sans cela, nous savons ce que nous devrions faire mais nous ne savons pas comment le faire[81]. La conclusion est que, bien que nous ayons des capacités naturelles, comme celle de raisonner sur les causes et les probabilités conditionnelles (une capacité qui est, sous une forme intuitive, limitée dès la petite enfance), nous pouvons continuer à commettre des erreurs dans l'usage des probabilités conditionnelles et dans le raisonnement causal même en présence de capacités plus sophistiquées et d'un certain nombre de connaissances. Le fait de penser de manière plus critique dépend beaucoup du contenu, ce qui est vrai aussi pour des formes plus spécialisées de pensée « experte », comme la pensée scientifique. Même dans le cas de la pensée scientifique, le succès dépend non seulement du fait de connaître les procédures, les stratégies, mais aussi de celui de savoir quand les appliquer et comment (Willingham 2007[82]). L'enseignement de l'esprit critique repose donc en partie sur le fait d'enseigner aux élèves des stratégies pour « mieux penser » et, en grande partie, sur la modalité et le moment de déployer ces stratégies.

De façon apparemment contre-intuitive, le développement de capacités d'EC dépend fortement de l'acquisition de connaissances de domaine. Transférer une compétence sur un contenu connu est en effet plus facile pour deux raisons : d'abord, parce que nous savons quel aspect prend la compétence en question dans le domaine considéré ; ensuite parce que les connaissances de domaine permettent plus facilement de reconnaître la structure profonde d'un problème et de dépasser le stade de l'examen superficiel (Willingham 2007[83]). Willingham cite des cas de réussite dans l'enseignement de l'EC, où les étudiants ont été exposés à des stratégies et ont été amenés à les appliquer de manière répétée. Mais il s'agit d'exemples internes à un certain domaine de contenu. Le scepticisme demeure lorsqu'il s'agit d'imaginer des programmes d'enseignement de l'EC visant à un transfert général, à l'application dans toutes les situations possibles de stratégies aussi générales qu'analyser, synthétiser, etc. car ces stratégies ne sont simplement pas les mêmes d'un contexte à un autre. Les objectifs de l'enseignement de l'EC, pour être réalistes, doivent donc être internes à un domaine. Rien ne peut suppléer l'expertise de domaine, lorsqu'il s'agit de penser de façon critique, de reconnaître la structure profonde d'un problème et de penser à adopter les stratégies appropriées (Willingham 2019[84]). À la fin, Willingham propose un plan en quatre étapes pour enseigner l'EC (Willingham 2019[85]) :

- identifier ce qui compte comme penser de façon critique dans chaque domaine : en histoire, nous ne prenons pas en compte les faits de la même manière qu'en sciences, par exemple ;
- enseigner les capacités et stratégies correspondantes de manière explicite ;

- identifier les contenus de domaine importants à maîtriser pour pouvoir penser dans le domaine en question ;
- planifier sur le long terme et re-exposer les apprenants aux capacités à apprendre, les faire pratiquer (pendant trois à cinq ans).

Caractères communs aux approches psychologiques et philosophiques

Une division relativement nette est tracée en littérature entre approches philosophiques et approches psychologiques : les premières sont plutôt normatives et les secondes plutôt descriptives (Lai 2011, Sternberg 1986).
Si les approches philosophiques indiquent des normes et standards à respecter pour bien penser, les approches psychologiques fournissent plutôt des listes de capacités cognitives à développer ou à exercer afin d'améliorer la prise de décision, rendre la pensée plus efficace ou tout simplement exercer des formes de pensée d'ordre supérieur qui s'opposent à d'autres formes plus simples, moins adaptées à des contextes complexes. En d'autres termes, pour les psychologues s'y étant intéressés, l'EC relève d'un ensemble de capacités utilisées pour atteindre un certain but ou remplir une certaine fonction, comme résoudre des problèmes, prendre des décisions, apprendre des nouveaux concepts. Ces capacités sont considérées comme nécessaires pour donner des réponses appropriées ou efficaces dans des situations complexes, là où des réponses automatiques, irréfléchies, basées sur l'application de formules ne peuvent pas donner le résultat espéré.

En réalité, cette distinction est moins nette : les approches philosophiques comprennent des listes de capacités (*skills, abilities*), exactement comme les approches psychologiques proposées par des auteurs comme Diane Halpern ou Robert Sternberg. Et les approches psychologiques comportent également des normes, comme les approches philosophiques : Richard Nisbett parle de l'importance des règles de la logique, des règles des mathématiques statistiques, des stratégies pour identifier les bonnes preuves ; Robert Stanovich oppose des processus de pensée rapides à des processus qui peuvent utiliser des règles apprises ; Diane Halpern indique la bonne manière d'utiliser les capacités naturelles pour arriver à une pensée plus efficace, et elle se réfère à des critères et normes. Ainsi, même si les psychologues n'ont pas développé une réflexion sur la notion de critères, les concepts de standards, de normes pour bien penser, de stratégies pour aider la pensée sont présents de manière plus ou moins implicite dans les approches psychologiques.

Le débat autour du rôle des connaissances de domaine est également transversal à la division entre approches philosophiques et approches psychologiques. Ainsi, Diane Halpern et Tim Van Gelder (par exemple, Halpern 2007, Van Gelder 2005) considèrent que l'EC est constitué de capacités générales. Pour Daniel Willingham, il s'agit d'un ensemble de capacités s'exerçant sur des contenus spécifiques, ce qui rend compliqué le transfert des acquis (Willingham 2007, 2019).

Du débat autour du rôle des connaissances dépend également celui autour de la transférabilité de l'EC (voir Lai 2011[86]). Les auteurs les plus optimistes sont ceux qui considèrent l'EC comme une capacité générale (nous venons de citer Halpern 2007, Van Gelder 2005 et nous avons cité Lipamn 1987 à titre d'exemple pour les approche philosophiques). D'autres ont plutôt tendance à voir l'EC comme lié à des contenus et donc plus difficile à transférer (c'est le cas de Willingham 2007, 2019 mais aussi de Bailin 2002 parmi les philosophes). Certains, enfin, considèrent que la spécificité de l'EC n'est pas une objection à sa transférabilité (par exemple, McPeck 1990). De ce débat découlent des divergences entre les différentes méthodes éducatives et de test de l'EC.

Cette raison de divergence s'ajoute au nombre et au type de capacités considérées comme fondamentales pour définir l'EC, ce qui a un impact sur les capacités enseignées et évaluées (Lai 2011[87]).

Les composantes suivantes sont aussi présentes aussi bien dans les approches philosophiques que dans les approches psychologiques de l'EC :

- l'objectif commun est d'arriver à des jugements ou des décisions plus satisfaisants (mais la nature de cette satisfaction varie d'un auteur à un autre) ;
- les deux camps s'accordent largement sur les capacités ou habilités générales à cultiver pour développer ou exercer l'EC, et notamment l'analyse, l'évaluation et le jugement d'arguments, assertions, preuves ; l'utilisation d'inférences et de formes de raisonnement inductif et déductif ; la référence à la prise de décision et à la résolution de problèmes (Lai 2011[88]) ;
- plus particulièrement, la plupart des auteurs associent la composante critique de la pensée critique à la capacité ou activité d'évaluation de l'information, d'arguments, d'affirmations ;
- aussi, le concept de métacognition est souvent – pas constamment – présent. Il peut être différemment interprété d'un auteur à un autre. Il s'agit toujours de métacognition explicite, d'autoréflexivité et de monitorage des performances ou des raisons utilisées pour se justifier (Lai 2011[89]) ;
- un autre point commun entre approches philosophiques et approches psychologiques concerne le pessimisme quant aux capacités de l'EC naturel (voir Lai 2011[90]). L'idée même de l'EC naturel n'est jamais évoquée : l'EC est pauvre dans nos sociétés, et ceci même dans le cas de sujets adultes éduqués. D'où la nécessité de formes d'éducation pensées pour développer l'EC (au niveau disciplinaire ou général), non limitées à la transmission de connaissances. Cette considération de départ entraîne une autre conséquence : les différents auteurs – y compris les psychologues – ne se sont pas attardés à rechercher et à décrire les bases naturelles de l'EC qui rendent possible son éducation, mais ils se sont concentrés soit sur les objectifs à atteindre (philosophes), soit sur les obstacles à dépasser (psychologues). La seule exception concerne Deanna Kuhn. Celle-ci focalise en effet son approche plus particulièrement sur une capacité : la métacognition. De plus, l'approche de Kuhn est développementale, ce qui signifie qu'elle retrace les différentes étapes du développement naturel de la métacognition pour donner des indications éducatives consistant à dépasser les limites naturelles de cette capacité et en développer une forme experte (voir Lai 2011[91]). Cependant, même dans le cas de Kuhn, nous pouvons extrapoler que le vrai EC est seulement celui considéré comme expert et que la métacognition naturelle en est une base imparfaite.

2.2 Des raisons d'insatisfaction par rapport aux approches existantes de l'EC

2.2.1 Des définitions trop vastes pour être opérationnelles

Nous avons indiqué au début de cette partie que la première raison d'insatisfaction dans la littérature existante sur l'EC concerne l'absence d'une définition consensuelle. L'apparent accord entre différents auteurs est essentiellement dû au fait que les définitions sont souvent

assez génériques et que les taxonomies de l'EC comprennent une grande variété de capacités et de dispositions. Ainsi, l'EC est souvent associé à des compétences et attitudes aussi variées que (voir Ennis 2016) :

- considérer de manière analytique, attentive, toute croyance ou opinion et chercher des raisons pour croire ; connaître les méthodes de l'analyse logique et rationnelle des arguments et savoir les appliquer ; adopter des critères explicites dans le cadre des opérations de la pensée et les utiliser pour guider le jugement dans une variété de contextes ; savoir conceptualiser ; être capable d'analyser, de synthétiser et d'évaluer concepts et informations ; savoir utiliser ces capacités pour guider l'action ; utiliser la raison pour prendre de meilleures décisions, raisonner ; savoir résoudre des problèmes ; posséder des capacités métacognitives, réflexives ; posséder des capacités de raisonnement algorithmique et logique ; faire preuve d'une clarté d'expression, de capacités d'argumentation ;
- se livrer à l'exercice volontaire de toutes les capacités citées ; avoir une disposition à la réflexivité, à l'exercice de la rationalité ; agir sur la seule base de raisons ; faire preuve d'une forme d'ouverture d'esprit, du désir d'être bien informé, de flexibilité ; faire preuve d'une certaine curiosité, d'une capacité à reconnaître un manque d'information, d'une capacité à changer d'idée et à suspendre son jugement ; être toujours prêt à se mettre en discussion ; faire preuve de propension à rechercher la raison, à questionner et à se questionner ;
- être sensible aux aspects éthiques et aux valeurs épistémiques, à l'équité ; respecter la vérité et respecter son interlocuteur.

Bien que convaincantes en premier lieu, les approches qui ont recours à ce type de taxonomies se révèlent insatisfaisantes, notamment dans une optique opérationnelle. En effet, si le concept d'EC venait à recouvrir l'ensemble des capacités et dispositions listées ci-dessus, il serait trop **vaste**. Associer la notion d'EC avec un ensemble trop vaste de capacités pose problème pour l'évaluation et l'éducation : il devient tout simplement impossible de traiter l'ensemble des dimensions liées à ce concept. Prenons à titre d'exemple la définition fournie par le panel Delphi (Facione 1990). Celle-ci établit que l'EC est un idéal qui consiste à émettre des jugements, de façon volontaire, réflexive, et à pouvoir justifier ces jugements sur la base de considérations de type méthodologique, conceptuelles ou des preuves dispo nibles. Pour y arriver, il est donc nécessaire de maîtriser l'ensemble des capacités (et sous-capacités) de l'EC, c'est-à-dire interpréter, analyser, évaluer, réaliser des inférences correctes, expliquer, s'autoréguler. Ces capacités incluent ensuite des sous-catégories : catégoriser, décoder la signification, clarifier le contenu, examiner les idées, identifier les arguments, les analyser, évaluer les affirmations et les arguments, chercher des preuves, imaginer des alternatives, arriver à des conclusions, présenter les résultats, justifier ses procédures, présenter ses arguments, s'auto-examiner, s'autocorriger. S'agirait-il, dans le cadre d'un programme d'éducation de l'EC, de traiter toutes ces capacités ? Et, dans ce cas, qu'est-ce que ce programme aurait-il de spécifique en relation avec l'EC, et de différent par rapport à un programme éducatif qui viserait à développer l'ensemble des capacités de pensée, de raisonnement, de résolution des problèmes, d'expression orale et écrite, aussi bien que les capacités métacognitive de ses candidats ?

S'il repose sur une variété trop importante de fonctions, et si la définition est trop générale, le concept d'EC perd de fait sa traitabilité. Des choix doivent donc être opérés pour réduire l'étendue des compétences à enseigner pour améliorer l'EC.

2.2.2 Des définitions trop vagues, qui peuvent donner lieu à des malentendus

La deuxième raison d'insatisfaction concerne la nature **vague** de certaines capacités et dispositions associées à l'EC, comme la capacité à « changer d'idée », à « faire attention aux faits » ou à « rester ouverts aux idées des autres ». Nous pouvons citer encore une fois le

rapport Delphi, qui décrit le penseur critique comme curieux, bien informé, confiant en la raison, ouvert, flexible, équilibré dans ses évaluations, honnête face à ses propres biais, prudent dans ses jugements, prêt à revenir en arrière, clair, ordonné, diligent dans la recherche d'informations, raisonnable dans la sélection des critères, focalisé sur l'investigation, etc. (Facione 1990).

Sans plus de précision, ces indications se prêtent à être mémorisées comme des concepts abstraits mais ne elles ne sont pas facilement déclinables de manière opérationnelle dans des situations bien précises. Faire attention aux faits ou déterminer la valeur des preuves par exemple sont des considérations difficiles à mettre en pratique si nous ne savons pas ce qu'est une preuve, voire une bonne preuve, dans le contexte donné (voir à ce propos la critique faite par Willingham 2007 aux approches éducatives de l'EC basées sur des principes généraux). S'il repose sur des notions vagues, le concept d'EC perd donc en **pertinence**.

Une définition trop vague peut en outre impliquer des conséquences opposées, voire indésirables. Le fait que l'EC soit associé à une vigilance envers les mauvaises informations et à un appel à plus d'ouverture d'esprit illustre de tels paradoxes. Des dispositions comme celle qui consiste à être prêt à revenir en arrière sur ses positions, si laissées sans spécifications, peuvent être interprétées de manière erronée. Changer systématiquement d'idée n'est pas plus désirable que ne jamais le faire. En effet, raisonnablement, nous ne sommes pas aussi prêts à abandonner nos idées quand celles-ci sont étayées par des preuves nombreuses et solides que quand elles sont au contraire basées sur des impressions ou de vagues intuitions. Il est important de préciser que le penseur critique n'est pas toujours à la recherche d'informations, mais qu'il sait s'arrêter sur une position en bonne confiance, si celle-ci est supportée par de bonnes raisons et preuves. Dans le cas contraire, le penseur critique se dépenserait totalement dans la mise en question constante de ses positions ou de celles des autres. La nature vague et abstraite de ces capacités et dispositions fait donc que la notion d'EC peut glisser facilement vers des formes indésirables de relativisme (douter de tout), de paralysie de l'action (suspendre le jugement, questionner tout) et vers des contradictions pratiques (rester ouvert, être méfiant). Les définitions existantes rendent parfois difficile de tracer une limite claire entre l'EC et les attitudes propres aux théories complotistes qui se caractérisent par une attitude relativiste et la construction de « mille-feuille argumentatifs » sans fin (Bronner 2013). S'il repose sur des notions vagues, le concept d'EC perd donc en **validité**.

2.2.3 Des définitions qui font référence à des concepts ambigus

Certaines définitions et taxonomies rapprochent l'EC de concepts tels que la raison, le raisonnement, la rationalité ou le raisonnable. C'est le cas par exemple de Willingham (2007) qui identifie l'EC avec trois catégories d'activités mentales : raisonner, émettre des jugements ou prendre des décisions et résoudre des problèmes. Ces concepts ne sont pas mieux spécifiés et définis que le concept d'EC lui-même. Il suffit de penser aux débats sur la définition du concept de raison et de ration alité (voir Mercier & Sperber 2011[92] et Mercier & Sperber 2017 pour une discussion à ce propos). Dans la littérature en psychologie du développement, le mot « raisonnement » est associé à des formes spécifiques de représentation et inférence qui concernent les objets, les agents, les quantités, la géométrie, les autres (Spelke & Kinzler 2007). Nous parlons également de raisonnement causal (Gopnik & Schulz 2007), de raisonnement par analogie (Gentner 1989), et nous en cherchons les origines précoces dans le développement cognitif de l'enfant. Il n'y aurait donc pas un seul type de raisonnement, indépendant de son contenu, ou de ses moyens.

Le plus souvent, le terme EC est cependant associé à une acception spécifique de raisonnement, développée dans le cadre de la psychologie du raisonnement et de la décision, notamment le modèle « dual » connu du grand public comme « système 1/système 2 » et le modèle associé du programme « heuristiques et biais » (c'est le cas notamment de Stanovich

& Stanovich 2010 et de Nisbett 2015). L'idée qui sous-tend le programme « heuristiques et biais » – par exemple telle qu'elle est exprimée dans le texte fondateur de Kahneman et Tversky (1974) – est la suivante : dans des conditions de jugement en situation d'incertitude, nous avons recours à des solutions-raccourcis qui sont en nombre limité, rapides à se mettre en place et simples. Ces heuristiques sont généralement utiles mais peuvent parfois nous induire en erreur. Quand une heuristique donne lieu à une erreur systématique, nous parlons de biais (Kahneman et al. 1982, Gilovich et al 2002, Kahneman & Tversky 1996). Ainsi, l'EC est associé au fonctionnement d'un « système 2 » qui serait rationnel, lent, algorithmique, souvent correct, et à l'idée d'inhibition ou de contrôle sur les biais cognitifs (résistance aux biais, *debiasing)*. Cette vision a donné lieu à débats et critiques, notamment dans le cadre d'une approche évolutionniste mettant l'accent sur la nature adaptative des processus cognitifs – par analogie avec les autres traits sélectionnés au cours de l'évolution (Cosmides & Tooby 1994, Gigerenzer 1991, 1996, 2004, Gigerenzer et al. 2008, Haselton et al. 2009). Gigerenzer, Tood & The ABC Research group (1999) ont par exemple opposé au programme des « heuristiques et biais » (solutions « *quick and dirty* ») le fait que certaines heuristiques permettent de donner des réponses rapides, économiques et correctes (« *fast and frugal* »). Ce cadre envisage les heuristiques comme faisant partie d'une « boîte à outils » mentale adaptative, sélectionnée au cours de notre histoire évolutive en réponse à des tâches et problèmes spécifiques qu'elles permettent de manière efficiente. Il serait donc erroné de trop souligner le caractère négatif ou biaisé de telles solutions. Le programme « heuristiques et biais » est aussi critiqué à cause de la nature artificielle des situations de test de laboratoire dans lesquelles les biais sont mis en évidence. Il suffirait de modifier ces situations pour voir les biais disparaître, ou du moins s'affaiblir. Par exemple, les études sur les biais cognitifs mettent en évidence les limites du raisonnement probabiliste. Mais la reformulation des problèmes posés en termes de fréquences permet d'augmenter le nombre de réponses correctes ; ceci parce que, en conditions naturelles, nous nous appuyons spontanément sur le constat qu'un événement est plus ou moins fréquent (Gigerenzer et al. 2008). Enfin, les approches évolutionnistes adaptatives reprochent à la vision « heuristiques et biais » d'avoir uniquement recours à des normes externes pour juger si un comportement est rationnel ou irrationnel. Dans l'approche « heuristiques et biais », la norme est établie sur la base de la désirabilité d'un certain type de comportement. Alors que, dans l'approche adaptative, nous cherchons à établir si la norme correspond réellement aux objectifs auxquels le système cognitif répond en vue d'augmenter la valeur sélective de l'individu, dans des conditions données. Ainsi, si la recherche de vérité est une norme importante du point de vue culturel, rien n'assure que les solutions sélectionnées au cours de l'évolution répondent uniquement à de telles normes.

Cette sélection s'est faite sous de multiples contraintes – dont la recherche de la vérité, mais pas seulement – et dans des environnements complexes, que nous négligeons souvent dans une approche « heuristiques et biais » classique. Sur la base de ces considérations, certaines tendances considérées comme menant à des prises de décisions irrationnelles – l'aversion au risque par exemple – peuvent être réinterprétées comme des stratégies très efficaces selon que nous considérons une optimalité plutôt qu'une autre (par exemple, minimiser la variance des gains – sauf dans les conditions où des gains forts sont nécessaires – plutôt que simplement maximiser les gains) (Cosmides 1989, Cosmides & Tooby 1996, Gigerenzer 1991).
L'approche évolutionniste peut ainsi révéler le caractère adaptatif de décisions *a priori* biaisées, et la ou les pression(s) évolutive(s) qui a (ont) façonné les processus de décision. La critique évolutionniste aux approches de type « heuristiques et biais » permet en outre de mettre en évidence la nécessité de questionner l'origine des erreurs et biais qui affectent la prise de décision et la formation d'opinions et de croyances. Ceci afin de cibler des stratégies d'amélioration réalistes plutôt que de simplement admettre ces biais et chercher à les contrer. Opposer de façon trop nette les deux approches « *quick and dirty* » and « *fast and frugal* » nous paraît cependant excessif. Les deux approches reconnaissent en effet le fait que le fonctionnement cognitif naturel nous met en condition de résoudre efficacement de nombreuses situations, mais qu'il présente également des limites, voire qu'il peut lui-même nous induire en erreur. L'idée de biais et de stratégies pour dépasser ou éviter certains biais

n'est donc pas à rejeter dans le cadre d'une approche de l'EC, mais il serait excessif – considération faite des critiques existantes – de réduire l'éducation de l'EC à une lutte contre les biais et l'EC lui-même à une modalité de pensée qui permettrait de se « débiaiser » (une vision présente chez des auteurs comme Keith Stanovich, voir Stanovich & Stanovich 2010).

2.2.4 Des définitions trop peu réalistes

Nous constatons que les approches existantes à l'EC font souvent preuve d'un certain **manque de réalisme psychologique**. La littérature philosophique en particulier fait souvent référence à une sorte de penseur idéal sans nécessairement s'interroger sur l'atteignabilité des objectifs qui sont posés pour ce penseur idéal, ni sur ses capacités « de départ » (Lai, 2011). À titre d'exemple, nous pouvons citer la définition qui a émergé du travail de recherche de consensus autour de l'EC (méthode Delphi) coordonné par Peter Facione (Facione 1990[93]).

La littérature en psychologie, de son côté, ne prend pas la peine de définir les bases naturelles de la pensée critique en termes développementales (exception faite du travail de Deanna Kuhn). L'effort descriptif est souvent porté sur l'identification des obstacles (les biais) qui s'interposent sur la voie de la réponse « correcte » ou du choix considéré *a priori* comme optimal, puis sur la proposition de stratégies de remédiation (c'est le cas de Richard Nisbett et de Keith Stanovich en particulier, voir Nisbett 2015 et Stanovich & Stanovich 2010). Il manque cependant une analyse de l'EC en termes de capacités cognitives le rendant en premier lieu possible, de leur fonction spécifique en lien avec l'EC, de leurs limites. Aucun des auteurs analysés, en dehors de Kuhn (voir Kuhn 1999), ne pose la question de l'EC en dehors du cadre éducatif. Pour les philosophes comme pour les psychologues rencontrés, l'EC est donc un objectif à atteindre, un état qui naît de stratégies artificielles ou du moins apprises. Nous avons vu que Kuhn identifie au contraire un groupe de capacités métacognitives comme étant la base naturelle de l'EC, et cherche ensuite à décrire le développement ontogénétique de ces capacités, avant toute forme d'éducation, jusqu'à ce qu'elles atteignent un plateau rendant nécessaire un effort éducatif pour être dépassé. Son approche du développement métacognitif de l'enfant est cependant datée. Le problème se pose donc de replacer l'EC dans le cadre actuel des connaissances sur la cognition.

Un aspect supplémentaire du manque de réalisme des approches courantes est représenté par l'absence de références à certains aspects « écologiques » de l'exercice de la pensée critique. Les aspects sociaux de la cognition ne sont ainsi que peu représentés (des exceptions étant représentées par Bailin & Battersby 2016 ; Bailin 2011 ; Kuhn 1999). Or, l'exercice de l'esprit critique a lieu dans un contexte souvent social. Nous échangeons nos informations avec d'autres, en recevons et en passons à notre tour. Les informations qui nous viennent d'autres personnes (informations de seconde main) nous arrivent avec des indices réputationnels, des considérations concernant le prestige de la source, son expertise présumée, sa familiarité avec nous – donc sa fiabilité (Origgi 2015). Les arguments que nous sommes amenés à juger existent dans un contexte culturel et font l'objet de prises de positions idéologiques, liées à des groupes d'opinion (Kahan 2015). Abstraire l'EC de ce contexte risque d'en faire un concept **débranché** de sa propre réalité.

Nous exerçons en outre nos capacités en relation avec d'autres. Nous discutons nos positions, nous argumentons. Nous ne nous limitons pas à mener nos analyses, évaluations, raisonnements dans une situation d'isolement, mais, souvent, dans une situation d'échange (voir Mercier et Sperber 2017, Mercier et Sperber 2011, Trouche et al. 2016 pour une approche argumentative du raisonnement, dans le sens où raisonner est un processus qui prend son sens dans l'échange d'arguments).

La dimension sociale de l'EC devient d'autant plus importante que ce concept est souvent évoqué en relation avec des phénomènes de nature culturelle et sociale tels que les complotismes, la circulation de fausses informations et leur influence éventuelle sur la prise de décision (*fake news*), les phénomènes de méfiance envers des savoirs établis sur des

preuves solides (la nature sûre de la vaccination, les faits à l'appui du réchauffement climatique) (voir Halpern 2013). Dans ce cadre se pose le problème de la confiance et à qui l'accorder (Oreskes 2019, Bronner 2013), donc de la capacité à évaluer les sources d'information, leur degré de fiabilité, leurs compétences, les éventuels leviers utilisés pour convaincre, mais aussi les facteurs qui rendent un contenu d'information plus attirant et facile à retenir qu'un autre. D'où viennent ces capacités ? En sommes-nous naturellement pourvus ? Comment accordons-nous notre confiance et pourquoi ?

C'est pour répondre à ce type de questions que nous considérons nécessaire de fonder la théorie de l'EC sur une analyse approfondie du fonctionnement cognitif naturel, de prendre en compte les aspects développementaux, les conditions écologiques dans lesquelles les capacités de l'EC s'expriment, leurs bases neurocognitives.

2.3 L'EC, qu'est-ce que c'est ?

Sur la base des considérations sur les limites des définitions et approches existantes, nous proposons de fournir une définition minimaliste de l'EC. Celle-ci permettra facilement d'isoler les capacités cognitives directement impliquées dans l'EC de celles accessoires. Nous pourrons plus facilement estimer où commence et où se termine une action visant à développer l'EC et identifier des outils d'évaluation mesurant spécifiquement l'EC.

En première approche, nous définissons l'EC comme *l'ensemble des capacités et des critères qui permettent d'évaluer la qualité épistémique des informations disponibles et de doser de façon conséquente notre confiance en ces informations, en vue de prendre une décision, de se forger une opinion, d'accepter ou de rejeter une affirmation à bon escient.*
Évaluer la qualité épistémique d'une information signifie se demander si l'information a de bonnes chances de correspondre à la réalité, donc si elle mérite notre confiance : est-elle plausible à la lumière des connaissances existantes ? Est-elle basée sur des preuves solides ? La source de l'information est-elle fiable ?

Bien que cette définition soit plus restreinte que la plupart de celles rencontrées, elle ne leur est pas étrangère. En effet, la notion d'évaluation est présente dans la plupart des approches rencontrées. Il suffit de citer le rapport Delphi, qui inclut l'évaluation parmi ses six composantes (Facione 1990) et les approches psychologiques de Halpern et Kuhn qui font explicitement référence à l'évaluation comme à la composante critique de l'esprit critique (Halpern 2013) et à l'évaluation de preuves à l'appui en tant que critère pour distinguer opinions et connaissances (Kuhn 1999). Une définition de l'EC basée sur la fonction d'évaluation de la qualité épistémique des informations présente donc le double avantage d'être spécifique et minimaliste et, en même temps, de s'insérer dans la continuité des approches existantes. Par conséquent, on ne pourra pas lui reprocher d'être tout simplement *une autre* définition de l'EC ou une définition d'*autre chose* que de l'EC.

Quelques distinctions se révèlent toutefois nécessaires, notamment afin d'identifier des mesures de l'EC :

- il importe de prendre en considération séparément les mécanismes de l'évaluation de l'information et ceux de la prise de décision. Lorsqu'un sujet prend une décision, il le fait sous l'influence d'un très grand nombre de contraintes, et sa capacité à évaluer de manière juste les informations n'est qu'un aspect du processus de décision. Les bénéfices liés à son intégration dans un groupe social par exemple peuvent prendre le pas sur l'acceptation de connaissances fiables (voir, à ce propos, la littérature sur le raisonnement motivé et les raisons socioculturelles qui sous-tendent des choix comme celui de la vaccination, ou l'acceptation de la théorie de l'évolution et du climat : Kahan et al. 2011, Kahan 2015, Kahan & Stanovich 2016, Kahan et al 2010). Il est

donc souhaitable de séparer ces deux aspects même si, dans la pratique, cela peut ne pas être simple à faire [la littérature existante ne fait pas nécessairement cette distinction. Ainsi, par exemple, (Halpern 2013) considère que l'EC coïncide avec les capacités qui permettent de prendre des décisions ou de résoudre des problèmes de manière plus efficace et plus efficiente] ;

- lorsque nous sommes aux prises avec une opinion ou à une décision, nous pouvons nous appuyer sur une évaluation plus ou moins approfondie des informations disponibles. Nous pouvons alors justifier notre choix sur la base d'arguments tirés de cette évaluation, enchaîner les arguments de manière cohérente, voire les présenter à un interlocuteur de manière éloquente. Le processus qui consiste à exposer et à expliquer nos choix sur la base d'arguments est cependant un processus à part par rapport à l'évaluation initiale de la qualité de l'information. Il ne rentre donc pas dans notre définition restreinte de l'EC, mais peut être considéré comme une capacité accessoire. Rappelons que, dans le cas de la définition fournie par le panel Delphi, l'EC comprend au contraire aussi bien les capacités d'évaluation que celles d'explication. De même, le panel Delphi a inclus dans sa définition les capacités d'interprétation, d'analyse et d'inférence (Facione 1990). Il s'agit dans ce cas de capacités associées à l'évaluation de la qualité d'une information, mais qui peuvent aussi rentrer dans d'autres processus de pensée et de raisonnement, et être séparées de l'évaluation elle-même. Les capacités d'inférence par exemple sont omniprésentes dans le fonctionnement cognitif (voir à ce propos Mercier & Sperber 2017). Les capacités d'interprétation sont souvent préalables (par exemple, l'interprétation d'un texte est préalable à l'évaluation de la qualité épistémique de son contenu mais les deux processus sont distincts). Celles d'analyse de la construction logique d'un argument sont complémentaires. Pour cette raison, contrairement au panel Delphi, nous n'incluons pas ces capacités dans notre définition spécifique de l'EC ;
- plusieurs auteurs font référence à la notion de métacognition ou d'autorégulation (par exemple, Kuhn 1999, Halpern 2013, Facione 1990, Ennis 2016). L'utilisation de cette notion n'est pas univoque dans la littérature sur l'EC et ailleurs. Kuhn par exemple l'utilise pour parler de métaconnaissance, de recours à des stratégies et aussi comme une forme d'épistémologie naïve (Kuhn 1999). De plus, la compréhension des capacités métacognitives et de leur développement a évolué récemment. Seule une analyse approfondie de la littérature récente en sciences cognitives pourrait donc nous dire dans quelle mesure l'EC en tant qu'évaluation de la confiance que nous pouvons mettre dans une certaine information dépend d'une forme ou d'une autre de métacognition.

Sur la base de la définition fournie, il est maintenant possible de se tourner vers la littérature en sciences cognitives afin d'identifier les bases naturelles de l'EC.

3. Les bases cognitives de l'EC

3.1 Naturaliser l'EC

À quelques exceptions près, les approches philosophiques et psychologiques tendent à ignorer la dimension « naturelle » et développementale de l'EC. L'EC est défini comme un objectif à atteindre, notamment grâce à une éducation dédiée, et les approches courantes portent donc l'attention plus sur ce qui lui fait obstacle et sur les méthodes permettant de le développer que sur ses éventuelles bases naturelles. Le message implicite qui émerge est que l'EC n'est pas « naturel », mais une conquête de l'éducation.

Souvent nous entendons d'ailleurs décrire nos concitoyens (et nous-mêmes) comme étant crédules, prêts à tomber dans les pièges des *fake news* et dans les mailles de la persuasion, sans aucune dotation naturelle pour s'en protéger. Bien sûr, nous pouvons tous proposer des anecdotes à l'appui de cette impression, et l'attention portée par les médias au phénomène des *fake news* ne fait que nourrir ce constat de fragilité face à la manipulation ou aux croyances douteuses. Mais sommes-nous réellement – par nature – complètement démunis d'EC ?

Nous allons rechercher dans la littérature concernée les réponses aux questions suivantes : existe-t-il des fonctions/processus cognitifs naturels qui nous rendent aptes à évaluer l'information et à calibrer notre confiance dans l'information ainsi évaluée, en vue d'une prise de décision ? Quel est le parcours développemental de ces fonctions/processus ? Quels sont leurs limites ? Plus généralement, la question qui guide cette deuxième partie est la suivante : existe-t-il des bases naturelles de l'EC ? Si oui, comment se développent-elles naturellement ?

La définition restreinte d'EC que nous avons proposée à la fin de la première partie nous permet de nous mettre plus facilement à la recherche des bases cognitives qui rendent l'EC possible.

Toutes ces questions pourraient paraître purement théoriques. Se limiterions-nous à chercher à mieux connaître l'EC en tant que produit de la cognition humaine ? En réalité, l'objectif de l'approche naturaliste présentée ici est opérationnel – comme pour les autres approches philosophiques et psychologiques à l'EC. Identifier les soubassements naturels de l'EC nous permet de donner une base de départ objective à l'éducation de l'EC. Identifier les bases naturelles de l'EC permet également d'établir de manière plus objective les limites de celui-ci, et ainsi de concevoir des actions pédagogiques ou des aménagements permettant de le dépasser.

L'« EC naturel » qui correspond à ces bases est en effet limité et ne nous sert pas toujours bien en toutes circonstances. Les accusations de crédulité ne sont pas complètement fausses, d'ailleurs, et force est de constater que nous ne sommes pas naturellement armés pour faire face aux défis contemporains, aux formes multiples de circulation de l'information, pour distinguer les connaissances les plus solidement fondées des opinions populaires largement propagées.

Il en va de même de notre capacité à construire des connaissances objectives. La littérature en psychologie du développement est riche d'exemples illustrant la présence d'une forme naturelle de raisonnement scientifique chez l'enfant. Le bébé raisonne déjà sur les relations causales, possède des conceptions et manifeste des attentes concernant les lois de la physique, distingue le monde du vivant, animé, du monde inanimé et attribue des états mentaux à autrui. Il explore en connaissance de cause, guidé par une curiosité qui n'est pas aléatoire mais rationnelle, sous certains points de vue du moins. Il possède un outillage lui permettant de

mettre à jour les hypothèses qu'il émet jour après jour grâce à son expérience et à ses connaissances préalables. Cependant, même s'il est si bien doté dès la naissance, ce bébé n'est pas un scientifique. Comprendre les explications de la science professionnelle, abandonner celles plus intuitives qui lui viennent de son bagage d'intuitions et d'expériences, est loin d'être simple. Même à l'âge adulte, l'être humain lutte pour obtenir une compréhension avancée de la réalité. Les experts dans le domaine de la construction des connaissances sur le monde naturel – les scientifiques de profession – sont pour cela dotés de toute une boîte à outils cognitifs afin de vérifier nos intuitions et de mettre sous contrôle certains de nos biais qui entravent une connaissance objective de la réalité. Les scientifiques peuvent ainsi arriver à des explications objectives, bien que souvent contre-intuitives. Les méthodes sont le fruit d'une longue histoire et de notre évolution culturelle. Au cours de son histoire, l'être humain a en effet produit des artefacts, des stratégies, afin d'outiller l'esprit scientifique naturel, déjà présent chez le bébé – notre dotation naturelle, forgée par des millions d'années d'évolution. Le résultat est une forme experte de connaissance du monde naturel et social.

Nous pouvons concevoir le développement de l'EC selon le même modèle que celui que nous venons d'adopter pour l'esprit scientifique. Le bébé vient au monde avec des capacités naturelles d'EC, qu'il exerce dans son quotidien pour faire face aux informations et départager des hypothèses en vue d'une prise de décision. Les tâches que l'enfant doit accomplir ne sont pas drastiquement différentes que celles auxquelles étaient confrontés nos ancêtres : les contraintes qui s'exerçaient alors ont façonné notre cognition, rendant le bébé apte à résoudre ces tâches, jadis comme maintenant.
Pour autant, il n'existe aucune raison de penser que notre EC est sans faille : d'une part, la sélection naturelle n'optimise pas tous les traits de manière simultanée, pas plus qu'elle n'optimise certains traits au-delà d'une valeur seuil (car cette optimisation ne se traduit plus en valeur adaptative) ; d'autre part, notre monde est aujourd'hui différent de celui auquel l'évolution nous a préparés. Il suffit de penser aux systèmes de circulation de l'information pour nous en rendre compte.

Nos exigences ont aussi changé : confrontés à un univers social et culturel en constante modification, nous cherchons probablement plus d'objectivité et de vérité que par le passé, ce qu'il ne serait pas naturel de faire si nous nous limitions à des considérations de type adaptatif immédiat. Nous sommes des êtres naturels et culturels, soumis à cette double évolution.

Ceci justifie que nous cherchions à développer artificiellement nos capacités naturelles d'EC. **Développer l'EC naturel signifie alors l'outiller de stratégies et de critères qui permettent de faire face à des situations différentes de celles pour lesquelles notre histoire évolutive nous a équipés.** L'EC naturel devient ainsi **EC avancé**. Les professionnels aux prises avec l'évaluation d'informations spécialisées – en médecine, en histoire, en physique... – ont besoin d'outils également spécifiques. Cela se traduit par différentes formes d'EC expert, capable de faire face aux défis posés par la nécessité d'acquérir des connaissances de plus en plus pointues et spécialisées.

À partir de ce cadre, nous analyserons la littérature en sciences cognitives (et plus particulièrement la psychologie du développement et les neurosciences cognitives) à la recherche des bases naturelles ou « *building blocks* » cognitifs de l'EC naturel. Nous terminerons par une excursion dans la psychologie du raisonnement et dans la psychologie évolutionnaire afin de tracer les limites de l'EC naturel. Enfin, nous fournirons des indications pratiques, opérationnelles, sur la manière – à partir de considérations préalables – dont nous pourrions nous y prendre, dans un contexte éducatif, pour développer l'EC vers des formes plus avancées.

3.2 Les bases cognitives de l'EC : la vigilance épistémique

Imaginons que nous devions décider où manger dans une ville inconnue. Au lieu de tester tous les restaurants, nous serions surement amenés à prendre des informations de seconde main. Par exemple, en demandant à quelqu'un quel est le meilleur restaurant du coin.

Cette tendance à acquérir des informations grâce aux autres est universelle et trouve sa raison d'être dans le gain de temps et d'énergie tout comme dans la diminution de prise de risque qu'elle représente, par comparaison à une prise d'information directe. Nous pourrions donner beaucoup d'exemples : suivre le conseil de quelqu'un qui nous dit de ne pas manger un certain type de champignon nous évite par exemple le risque de nous empoisonner. Cependant, pour que l'information fournie soit utile, nous devons nous assurer que notre informateur ne nous ment pas volontairement et qu'il en sait effectivement plus que nous.

En effet, la dépendance aux autres pour acquérir des informations vitales comporte le risque majeur de se laisser désinformer. L'enjeu est tel (pouvoir profiter de l'information de seconde main sans prendre de risques excessifs) qu'il semble plausible, d'un point de vue évolutif, que la pression sélective exercée sur nos fonctions cognitives ait donné lieu à des mécanismes capables de filtrer l'information et de nous protéger de la désinformation – même imparfaitement. Et ceci, par des moyens efficients en matière de coûts, et au moins approximativement corrects (Harris & Corriveau 2011, Sperber et al. 2010[94]). Tout cela ne pourrait être qu'une spéculation évolutionnaire – une « *just so story* » séduisante mais non vérifiable. Or, ces mécanismes supposés « logiques » d'un point de vue évolutif semblent en effet à l'œuvre dans notre fonctionnement cognitif, et ceci dès l'enfance.

Les études empiriques montrent en effet que nous sommes dotés d'un certain nombre de mécanismes, qui ont été étudiés notamment chez l'enfant dans une perspective développementale (Sperber et al. 2010, Harris, 2012) et réunis sous le terme de « vigilance épistémique » (Sperber et al. 2010[95]). Ces mécanismes nous aident à juger aussi bien la pertinence que la qualité « épistémique » de l'information de seconde main – et donc de conclure si les informations fournies par d'autres ont des chances de correspondre à la réalité (ou si notre informateur nous livre volontairement ou involontairement des informations incorrectes). Ainsi, nous pouvons rapidement déterminer si les informations fournies par d'autres sont susceptibles de correspondre à la réalité ou non. Ces processus d'évaluation ne sont pas nécessairement conscients, réfléchis ou coûteux, et ils peuvent se baser sur de simples indices. Le résultat de leur action est un état de confiance plus ou moins fort envers l'information donnée, qui se traduit d'un point de vue comportemental par une plus ou moins grande disponibilité à utiliser celle-ci pour se forger une opinion/croyance ou prendre un autre type de décision. Ces processus ont été étudiés, notamment chez l'enfant, dans une perspective développementale (Harris, 2012).

3.2.1 Les mécanismes de la vigilance épistémique

La vigilance épistémique appliquée à des informations de seconde main s'exerce de fait à deux niveaux : elle a pour objet à la fois les contenus – quoi croire – et les sources de l'information – qui croire.

Le premier groupe de mécanismes concerne les contenus de l'information transmise, indépendamment de sa source. Ils consistent par exemple dans l'évaluation de la cohérence avec nos connaissances préalables (*background beliefs*) ou avec d'autres constats que nous pouvons réaliser directement (et dans lesquels nous avons confiance). Il s'agit donc de mécanismes permettant d'évaluer la plausibilité du contenu (*plausibility check*) et sa pertinence (*relevance*)[96].

Le deuxième groupe de mécanismes s'applique à la source de l'information (*trustworthiness*). Celle-ci peut être porteuse de désinformation pour deux raisons distinctes : elle est en possession de bonnes informations mais veut nous tromper (malveillance) ou elle est bienveillante mais ignorante (ignorance). Des mécanismes traitent donc les deux possibilités[97].

L'idée de vigilance épistémique repose, outre sur des considérations d'ordre évolutionnaire, sur des constats réalisés notamment sur les jeunes enfants. Des études montrent que les jeunes enfants font preuve d'une confiance sélective (*selective trust*) : ils utilisent très précocement des critères permettant de choisir un informateur parmi d'autres de façon non aléatoire. Ils montrent généralement une préférence pour des adultes bienveillants et compétents[98].

Un exemple permet de comprendre l'articulation de ces mécanismes. Si nous demandons à un passant le chemin pour le bureau de vote de notre quartier, notre attention sera éveillée par le fait qu'il porte un signe distinctif de la Ligue des abstentionnistes ou qu'il semble improviser (évaluation de la source : malveillante ou ignorante) ou par le fait qu'il mentionne un chemin incluant des détours de plusieurs kilomètres (évaluation du contenu : non cohérent avec ce que nous connaissons du maillage des bureaux de vote ou de la géographie du quartier).

3.2.1.1 Les mécanismes de la confiance sélective chez l'enfant (vigilance dirigée sur les sources)

Un nombre significatif d'études montre que les enfants utilisent dès trois ans des critères (implicites) afin d'opter pour un informateur parmi d'autres de façon non aléatoire (confiance sélective, Harris, 2012) : d'un côté, des critères qui leur permettent de sélectionner un informateur après qu'ils aient pu interagir avec lui (familiarité de l'informateur, indices qui permettent d'inférer la compétence passée de celui-ci) ; de l'autre, des critères afin de choisir un informateur même quand ils se trouvent pour la première fois en contact avec lui (compétences génériques observées de manière directe ou *via* des indices de prestige ou de consensualité, indices de bienveillance, comme l'appartenance au même groupe linguistique) (Harris & Corriveau 2011).

Un premier groupe d'études concerne des enfants de maternelle. Dans une expérience-type, les jeunes enfants regardent deux vidéos. Dans chaque vidéo, un adulte se réfère à des objets ou à leurs fonctions. Les deux adultes fournissent des noms différents pour l'un de ces nouveaux objets. Par exemple, l'adulte A appelle l'objet un « *snegg* », l'autre un « *hoog* ». On demande ensuite à l'enfant : quel est le nom de l'objet « inconnu » ? L'adulte A est un adulte familier à l'enfant (une maîtresse de son école) et l'adulte B une maîtresse d'une autre école, inconnue de l'enfant. Résultat : les enfants de trois à cinq ans montrent une préférence pour l'adulte familier, A s'ils appartiennent à l'école A, B s'ils appartiennent à l'école B. Conclusion : les enfants n'ont pas choisi leur informateur au hasard mais ont eu recours à un indice, celui de la familiarité, pour opérer leur choix. Mais la familiarité n'est pas tout.

Dans une autre expérience, conduite avec des enfants de trois à quatre ans, les deux informateurs (non familiers des enfants) doivent d'abord nommer des objets communs, connus par les enfants. L'un des deux adultes se trompe en nommant ces objets (par exemple, il appelle « balle » une tasse). Si les enfants doivent ensuite demander le nom d'un objet nouveau et inconnu, ils vont décider de se tourner vers l'adulte qui ne s'est pas trompé sur les objets connus ; si les deux adultes fournissent des noms différents pour le même objet inconnu, les enfants choisissent le nom fourni par l'adulte le mieux informé. Ils orientent donc leurs choix sur la base des compétences montrées par l'adulte, et suivent l'adulte le plus compétent.

Une troisième expérience compare les effets de la familiarité et ceux de la justesse. Dans ce cas, l'un des deux informateurs est familier, l'autre non ; celui qui se trompe sur les objets communs est, pour un groupe d'enfants, l'informateur familier et pour l'autre groupe

l'informateur non familier. Pour les enfants de trois ans, le fait que l'adulte familier commette des erreurs n'est pas suffisant pour lui retirer leur confiance. Mais, pour les enfants de cinq ans, le fait d'être correct devient une condition de confiance (*trust*) plus importante que la familiarité. Non seulement les enfants sont sensibles à la compétence spécifique de l'adulte qui sait nommer correctement des objets communs, mais ils tendent aussi à préférer celui qui montre des compétences génériques, comme celle d'utiliser une grammaire correcte (en présence d'un adulte qui prononce « *a shoes* » et un qui dit correctement « *some shoes* », les enfants préfèrent le deuxième comme informateur).

Un quatrième groupe d'expériences montre que les jeunes enfants sont sensibles à des indices d'ordre social, au moment de choisir leur informateur. Ils semblent par exemple préférer un adulte qui a un accent natif plutôt qu'un adulte avec un accent étranger[99].
Dans un cinquième groupe d'expériences, les enfants sont expérimentalement confrontés à un cas de désaccord entre deux adultes, l'un recevant l'approbation d'autres adultes, l'autre non. L'adulte préféré est celui qui reçoit l'approbation des autres.

Grâce à cette batterie d'études, nous pouvons montrer que les enfants sont sélectivement attirés par les adultes qui se montrent « culturellement conformistes », car ils reçoivent l'approbation du plus grand nombre[100]. Morgan, Laland & Harris (2015) ont testé l'influence du degré de consensus : ils ont présenté à des enfants d'âge compris entre trois et sept ans des images avec un nombre de points variable, et leur ont demandé d'estimer laquelle de deux images comportait le plus de points. La tâche variait en difficulté en raison du ratio de points entre les deux images. Les enfants de trois ans étaiet influencés par le consensus social seulement si celui-ci était total ; dans le cas contraire, l'influence du contexte social était aléatoire. À partir de sept ans, les enfants répondaient également à des majorités non absolues et de façon proportionnelle au niveau de consensus.

En conclusion, les enfants possèdent deux grandes catégories d'heuristiques leur permettant de jauger la fiabilité de leurs sources d'information : d'un côté, des heuristiques qui servent à sélectionner un informateur après qu'ils ont pu interagir avec lui (familiarité, compétence passée) ; de l'autre, des heuristiques permettant de choisir même quand ils se trouvent pour la première fois face à l'informateur (compétences génériques, appartenance linguistique, consensualité[101]).

Les critères utilisés par les enfants changent au cours de leur développement. Certaines trajectoires évolutives ont été mises en évidence, comme l'influence de plus en plus importante du critère d'expertise par rapport à la familiarité, ou la sensibilité à des formes de consensus plus mitigées et la capacité d'adapter sa confiance selon le degré de consensus perçu. Cependant, l'étude du développement des mécanismes et critères de confiance sélective au cours de l'enfance est encore un « *work in progress* ».

3.2.1.2 L'épistémologie naïve de l'enfant (vigilance dirigée sur les contenus)

Une vaste tradition attribue aux enfants la capacité de distinguer entre croyances (des représentations qui peuvent être fausses) et réalité seulement à partir de l'âge de quatre à cinq ans (Wellmann, Cross, Watson 2001). Récemment, ce dogme a été ébranlé par des études montrant que cette capacité est maîtrisée de manière implicite, non propositionnelle, déjà autour d'un an et demi à deux ans (Scott & Baillargeon 2017). Une solution proposée à cette apparente contradiction est celle de distinguer entre formes implicites et formes explicites d'attribution de croyances et d'autres états mentaux à autrui (Rakoczy 2017 ; Apperly & Butterfill 2009). Dans cette perspective, Mascaro & Morin (2015) ont étudié la capacité des enfants à traiter la notion de fausseté avant l'âge de cinq ans. À travers leur analyse de la littérature et avec plusieurs d'expériences, ils ont montré qu'elle se développe progressivement et que les enfants de deux ans ont déjà la capacité de considérer une affirmation comme fausse. Cependant, avant l'âge de quatre à cinq ans, ils ont du mal à

mémoriser des assertions fausses, et ils tendent à les corriger involontairement. Une interprétation possible pour ce type d'incohérence est que, bien que capables de comprendre la distinction entre vrai et faux, les enfants plus jeunes sont plus confiants dans leurs informateurs et tendent à leur attribuer des croyances vraies, donc à les corriger implicitement si ces informateurs manifestent des croyances clairement fausses.

Les enfants font donc preuve d'une forme d'épistémologie naïve précoce qui leur permet d'évaluer les contenus d'information comme étant vrais ou faux[102].
La capacité d'apprécier la valeur des preuves à l'appui d'une assertion a, elle, été étudiée plus dans le cadre de recherches sur le raisonnement causal et scientifique des enfants et des adolescents (Zimmerman 2000 ; Morris et al. 2012) que dans celui d'études concernant la vigilance épistémique.

Gopnik et al. (2001), Schulz & Gopnik (2004) et Schulz, Gopnik & Glymour (2007) ont montré comment les enfants entre deux et quatre ans utilisent l'observation de la réalité pour en inférer des relations de type causal dans des situations relativement complexes. (Avant cet âge, entre six et six-huit mois, les enfants sont déjà capables d'inférences causales sur la base d'indices simples comme l'ordre temporel et la contiguïté entre événements.)

Les enfants de six ans sont capables de distinguer un test discriminant d'un test non discriminant à partir d'une hypothèse et quand les deux tests leur sont proposés. Ils ne sont cependant pas encore capables d'imaginer par eux-mêmes un test discriminant.
Entre huit et douze ans, leur capacité à comprendre ce qui constitue une bonne preuve (*evidence*) de causalité est encore limitée : par exemple, ils ne savent pas systématiquement produire des tests sans facteurs confondants (si nous leur demandons de chercher la cause d'un effet, ils font varier plusieurs facteurs en même temps). De plus, ils ne prennent pas de notes de façon spontanée afin de structurer leurs observations (Klahr, Fay, & Dunbar 1993 ; Kuhn et al. 1995 ; Schauble 1990, 1996 ; Zimmerman, Raghavan, & Sartoris 2003). Concernant la compréhension de la notion de preuve, les enfants peuvent confondre leurs croyances initiales avec les données d'observation lorsque nous leur demandons de justifier une assertion. Ils démontrent en cela qu'ils ne maîtrisent pas la différence entre une connaissance préalable et des données d'observation (Schauble 1990).

Kuhn (1989, 2005, 2011) a étudié le développement des capacités métacognitives des enfants en relation, d'un côté, avec le raisonnement scientifique et, de l'autre, avec le développement de l'EC come capacité à raisonner sur ce qui constitue une connaissance par rapport à une simple opinion. Kuhn s'est intéressée en particulier à la capacité à articuler théories et preuves empiriques, donc hypothèses ou connaissances initiales et données d'observation. Selon elle, cette capacité implique trois sous-composantes : la capacité de représenter séparément théorie et preuve ; celle de traiter les théories comme des formes de représentations et non comme un état des choses du monde ; la reconnaissance qu'une théorie peut être fausse et que, pour déterminer si elle vraie ou fausse, il est nécessaire de se tourner vers les preuves disponibles. La considération la plus importante à laquelle arrive Kuhn est que les stratégies pour coordonner théories et preuves ne se développent pas sans instruction formelle. Il ne s'agit donc pas d'un développement progressif, par étapes, selon une trajectoire naturelle, mais d'un apprentissage piloté[103].

Nous avons donc affaire au domaine de la vigilance épistémique qui n'est pas complètement naturel, mais requiert une intégration culturelle, grâce à l'éducation. Cependant, les études citées s'intéressent plus à la capacité de l'enfant de produire qu'à celle de comprendre la valeur d'une preuve à l'appui d'une affirmation. Or, comme le montre l'étude citée concernant la compréhension d'un test discriminant (Schauble 1990), la capacité de reconnaître une bonne preuve pourrait précéder celle d'en produire. Dans l'état actuel des connaissances, d'autres recherches semblent donc nécessaires afin de pouvoir se prononcer sur les bases naturelles de la vigilance épistémique relative aux contenus.

3.2.2 Les mécanismes de la vigilance épistémique ne sont pas infaillibles

La théorie de la vigilance épistémique insiste sur l'idée selon laquelle les critères naturels que nous utilisons pour évaluer la qualité de l'information fournie sont plus ou moins sophistiqués et exigeants en termes cognitifs. En réalité, si nous voulions être très sélectifs, nous devrions mettre en place des critères plus coûteux, qui évalueraient notre confiance en s'adaptant aux circonstances : par exemple en reconnaissant qu'une même personne peut être compétente ou ignorante en fonction du domaine. Cependant, nous ne sommes pas toujours capables ou disposés à payer le coût de ces critères, ce qui fait que nous en utilisons d'autres plus superficiels et ayant une marge d'erreur plus importante[104]. Ainsi, nous pouvons nous baser sur le « caractère » de la personne et non sur ses performances dans la circonstance, voire sur son aspect physique comme indicateurs indirects de fiabilité. De façon plus générale, nous nous fondons sur des signes « non spécifiques ». Le risque existe donc que ces critères à bas coût ne soient pas assez raffinés pour certaines situations où un coût plus élevé peut être payé pour se garantir contre manipulation et incompétence[105].

Les performances des processus de la vigilance épistémique naturelle se dégradent, notamment dans des situations complexes ou inhabituelles, dans le cas d'un contexte social élargi – par rapport à celui ayant vu évoluer nos capacités – par le truchement des réseaux sociaux par exemple. Ainsi, nous avons tendance à juger de la qualité épistémique d'une information si elle est largement partagée. Si chaque membre du groupe arrive seul à une conclusion, faire confiance à l'opinion générale relève de la rationalité. Mais s'il s'agit d'une idée diffusée dans le groupe par la communication, elle peut largement être partagée pour d'autres raisons que son bien-fondé. Cependant, à cause de la situation décrite auparavant, l'indice « acceptée par beaucoup de monde » peut être interprété comme un signal positif concernant la qualité épistémique de l'information. Ainsi, les mécanismes mêmes de vigilance épistémique naturels peuvent devenir des causes d'erreur de jugement, à cause du recours à des critères faible et à bas coût[106].

3.2.3 Des formes sociales et culturelles de vigilance épistémique permettent de remédier aux limites de la vigilance épistémique naturelle. Naissance de la vigilance épistémique étendue

Face aux limites de la vigilance épistémique individuelle, des formes sociales et culturelles de cette même vigilance ont historiquement émergé. Il s'agit d'organismes institutionnels formés d'experts proposant des moyens pour une meilleure évaluation des informations qui circulent au sein d'un groupe. Sont considérés comme experts ceux qui possèdent des critères d'évaluation de l'information plus fins et adaptés que les critères spontanément utilisés par les non-experts. Comme dans le cas individuel, la vigilance collective doit être exercée à la fois par rapport au contenu et par rapport à la source. La science, avec ses mécanismes d'évaluation des preuves et ses structures institutionnelles de validation des sources (*peer review*), représente un cas exemplaire de vigilance épistémique sociale.
Ainsi, nous pouvons soutenir que la vigilance épistémique ne se limite pas à un exercice individuel. Comme dans le cas individuel, la vigilance doit dans ce cas être exercée par rapport à la fois au contenu et à la source. La science, avec ses mécanismes d'évaluation des preuves et ses structures institutionnelles de validation des sources (système de relecture des articles par les pairs – *peer reviewing*, réputation, revues, etc.), est un cas exemplaire de vigilance épistémique « sociale et distribuée[107] ».

Nous pouvons alors considérer que l'articulation entre les mécanismes individuels (psychologiques) de vigilance épistémique, les artefacts cognitifs (permettant de rendre

l'évaluation des informations plus contrainte, comme la méthode scientifique) et les dispositifs institutionnels d'évaluation des sources (le monde académique, les journaux, le *peer reviewing*) permet une forme étendue de vigilance épistémique, une forme de cognition distribuée au service de la vigilance épistémique ou de l'esprit critique – selon un modèle propre à la cognition distribuée (Hutchins 1995a, 1995b[108]).

3.2.4 Expliquer la crédulité, les opinions irrationnelles et les effets de persuasion

L'existence de mécanismes naturels de vigilance épistémique et l'identification de leurs limites entraînent des conséquences importantes pour toute théorie de l'EC.

La première est une conséquence optimiste : nous ne sommes pas des « idiots critiques » par nature. Nous sommes au contraire dotés de mécanismes de protection, qui reposent sur la sélection des informations et de leurs sources. Ces mécanismes agissent sur la base de critères à bas coût, sont faillibles et sont mis particulièrement en difficulté en cas de circulation rapide de l'information à distance. Mais ils constituent des bases naturelles pour notre EC face aux opinions d'autrui.

La deuxième conséquence se décline plutôt en termes d'interrogations : comment expliquer – à la lumière des mécanismes de la vigilance épistémique – la diffusion des *fake news* et des opinions « bizarres » que nous entretenons parfois ? Le rejet de positions scientifiques accréditées ? Tout cela sert-il uniquement à attribuer des limites à notre vigilance épistémique naturelle ? Autrement dit, quel rôle joue dans ces manifestations la manipulation de l'opinion opérée par les médias, les réseaux sociaux ou par des personnalités particulièrement influentes ? Serions-nous prêts à accepter les thèses les plus farfelues seulement parce que quelqu'un nous convainc de le faire ?

3.2.4.1 Nous ne sommes pas aveuglément crédules, mais plutôt conservateurs

Diverses expériences de psychologie sociale, conduites dans les années 1950-1960 par Solomon Asch et Stanley Milgram, respectivement sur l'influence du consensus et sur la déférence à l'autorité ont laissé une image très négative concernant notre capacité à résister aux influences d'autrui (Asch 1951 ; Milgram 1974).

Une vision largement répandue voit ainsi l'être humain comme « *wired not to seek truth but consensus* » : prêt à accepter aveuglément les thèses les plus folles parce qu'elles seraient largement acceptées ou portées par des sources prestigieuses (voir Mercier 2017 pour une analyse de cette littérature[109]).

La théorie de la vigilance épistémique rejette cependant l'idée d'une crédulité massive. Pour Mercier (2017, 2020), cette idée est d'autant plus irréaliste que notre mode de fonctionnement « par défaut » est plutôt un mode conservateur. Confrontés à des informations qui contredisent nos croyances préalables, nous nous engageons dans un contrôle de plausibilité. Si les nouvelles informations viennent de nous-mêmes (par exemple, nous observons quelqu'un dans le jardin alors que nous pensions qu'il était ailleurs), nous sommes prêts à changer de position et à mettre à jour nos croyances. Mais si l'information est communiquée par d'autres et si nous n'avons pas de bonnes raisons de croire en l'expertise ou en la bonne fois de cet informateur, nous tendons à camper sur nos positions. Le rejet des positions communiquées par d'autres serait d'autant plus radical que les positions initiales et l'information reçue divergeraient (Mercier 2017[110], Yaniv 2004[111]). Yaniv (2004) a ainsi testé notre susceptibilité d'accepter l'avis d'autrui quand nous avons déjà formé notre avis sur une question. Même si les participants sont motivés pour donner la bonne réponse (ils reçoivent un bonus pour les réponses correctes), les résultats indiquent que les sujets ont une attitude

biaisée envers leur propre avis. Plus les avis sont éloignés, moins les participants sont disposés à changer de position.

Ces considérations laissent penser que les accusations de crédulité qui sont souvent portées et attribuées à l'influence néfaste des autres (des médias, de la télévision, des campagnes de persuasion de masse…) pourraient être mal placées. Une analyse de la littérature effectuée par Mercier (2017) confirme que les campagnes politiques et religieuses ont moins de succès que ce que nous imaginons, ou du moins que leur impact est modulé par les contenus proposés.

Néanmoins, il semble important, dans le cadre d'une théorie de l'EC, d'approfondir rapidement la question de l'influence d'autrui. Dans les paragraphes suivants, nous collecterons donc des éléments pour répondre du moins partiellement à la question suivante : quand acceptons-nous de changer d'avis et quand acceptons-nous une information de seconde main ? Ou encore : quand nous laissons-nous convaincre ?

3.2.4.2 De bonnes raisons pour accepter les idées d'autrui

En présence de conflits d'opinions, même si nous conservons un biais positif en faveur de nos propres opinions, nous nous engageons tout de même dans l'évaluation de l'information d'autrui avec tous les mécanismes disponibles de la vigilance épistémique, à la fois ceux du contrôle sur les contenus (plausibilité, pertinence) que ceux de l'évaluation des sources (expertise et bienveillance). Nous prenons en compte les compétences de l'informateur, ses éventuels conflits d'intérêt, sa réputation morale, son prestige ; nous faisons, comme les enfants, attention à sa langue ou à son accent, donc à son appartenance au même groupe que nous, et à l'approbation qu'il reçoit par le groupe (Mercier 2017[112]). Ces propriétés pèsent dans notre propension à changer d'avis et à nous laisser convaincre.

Mercier (2017) insiste également sur le fait que, pour changer d'opinion, de bons arguments doivent nous être apportés en faveur de celle-ci. Il prend l'exemple d'un problème classique de raisonnement, celui de la batte et de la balle, que nous présentons ici dans une version différente.

La bouteille et son bouchon coûtent ensemble 1 euro et 10 centimes. La bouteille coûte 1 euro de plus que le bouchon. Quel est le coût du bouchon ?

Ce problème est souvent utilisé comme test de raisonnement et pour montrer l'existence de biais irrésistibles qui nous rendent irrationnels (Frederick 2005). Un pourcentage très élevé de participants ne parvient pas à donner la réponse correcte (environ 85 %, selon le dispositif expérimental utilisé, répondent que le bouchon coûte 10 centimes). Mercier montre, sur la base d'expériences menées en ligne, que le problème est plus facilement résolu en groupe. En effet, dans une situation de groupe, ceux qui ont la mauvaise réponse tendent à être convaincus par ceux qui ont la bonne, en particulier si elle leur est présentée avec des arguments solides (Trouche, Sander & Mercier 2014 ; Trouche, Shao & Mercier 2019 ; Mercier 2017[113]).

Notre propension à suivre l'avis d'autrui serait donc modulée par les mécanismes de la vigilance épistémique déjà mis en évidence (avec leurs limites) et par l'existence de bonnes raisons pour le faire.

3.2.4.3 L'influence des contenus sur l'acceptation des idées

Cependant, nous ne nous basons pas exclusivement sur des bonnes raisons et sur des bons arguments pour suivre l'avis d'autrui. Nous avons cité l'analyse de la littérature effectuée par (Mercier 2017) selon laquelle l'influence de politiciens, religieux et publicitaires sur nos opinions serait moins importante que nous nous imaginons. Cette influence existe néanmoins, modulée par les contenus promus[114]. Ainsi, nous pouvons nous laisser convaincre que nos intuitions ne sont pas en conflit, mais plutôt en accord, avec les idées promues par autrui, qu'elles soient bonnes ou mauvaises. (Un exemple de la façon dont nos propres intuitions peuvent nous induire à accepter des idées fausses est représenté par la pratique de la saignée : totalement inefficace, dangereuse, et néanmoins largement répandue. Certains attribuent son succès aux enseignements de Galien, mais était est en réalité enracinée dans nos intuitions de pureté[115].)

L'idée que certaines caractéristiques d'une idée la rendent plus attractive que d'autres remonte à une longue tradition : à l'idée de *meme* rendue populaire par Dawkins (2016, 1982[116]), à la théorie de l'épidémiologie des représentations développée par Cavalli-Sforza & Feldman (1981) et par Sperber (1996) et à la théorie de la transmission culturelle proposée par Boyd & Richerson (2005), Richerson & Boyd (2008). Les trois présentent l'application d'une approche darwiniste de l'évolution de la culture, *via* l'idée selon laquelle l'information culturelle se diffuse dans une population *via* ses interactions et la production d'artefacts sociaux[117]. Dans le cadre de cette tradition, l'influence des émotions – et notamment des émotions fortes (Heath, Bell & Sternberg 2001), négatives (Bebbington et al. 2015) et associées à un danger Blaine & Boyer (2018) – a été utilisée pour expliquer le succès de légendes urbaines (Heath, Bell & Sternberg 2001), croyances religieuses, rumeurs (Blaine & Boyer 2017[118]). Plus récemment, Acerbi (2019a, b) a analysé des sites connus pour publier des *fake news*. Il a ainsi mis en évidence que les informations mises en avant présentaient des spécificités par rapport à celles considérées comme moins attirantes. Notamment, les *fake news* à succès faisaient plus souvent appel à des émotions négatives liées à une menace, contenaient des éléments sexuels et faisaient appel aux réactions de dégoût[119].

Ainsi, selon ces exemples, nos systèmes de protection de type vigilance épistémique ne sont pas sans brèche et sans imperfection (Boyer 2018[120]). Non seulement ces systèmes sont limités[121], mais les caractéristiques de certains contenus d'information créent une brèche dans nos systèmes de défense parce que leur contenu est particulièrement saillant, attractif, ou pertinent d'un point de vue adaptatif[122]. Ces contenus rendent donc encore plus difficile d'exercer sa vigilance.

À ces considérations, il est nécessaire d'ajouter le fait que l'évolution de nos systèmes de vigilance pousse au développement d'armes encore plus puissantes de persuasion et de manipulation. Puisque tromper est avantageux, les systèmes pour tromper évoluent en même temps que ceux permettant de parer les tromperies. Le design de ces systèmes pourrait donc, du moins pour certains contenus et dans certaines conditions – ou à un moment précis de la course aux armements – donner un avantage à des stratégies de tromperie[123].

3.2.4.4 Le rôle de l'appartenance au groupe

D'autres raisons de la réussite des idées à succès peuvent paraître moins immédiates. Elles concernent notamment la capacité de certaines idées à cimenter des alliances. Selon Atran (2002) et Boyer (2018), certaines d'entre elles se prêtent particulièrement à la création de groupes avec un très fort niveau d'engagement. Il s'agit notamment des idées contre-intuitives, comme l'existence des fantômes, des miracles, etc. Adhérer à des idées contre-intuitives constituerait un signal fort d'appartenance à un groupe. L'adhésion même à ces idées serait donc motivée – de façon implicite, non assumée – par la volonté d'être à tout prix membres du groupe. De même, certaines rumeurs seraient en réalité utilisées comme des agrégateurs de consensus. Celui qui passe l'information n'a pas (du moins pas uniquement) la motivation d'informer mais (surtout) celle de créer une coalition pour une action commune

(Boyer 2018[124]). Ce type de mécanisme pourrait en partie expliquer le succès de théories du complot, des croyances religieuses les plus contre-intuitives, des *fake news* les plus invraisemblables.

Un courant d'études en psychologie sociale a développé une théorie cognitive et culturelle afin d'expliquer certaines positions anti-scientifiques – notamment des positions anti-vaccins et des positions sceptiques envers le changement climatique (Kahan 2015). Cette théorie permet d'expliquer la résistance aux explications d'ordre cognitif lorsque l'adhésion à des idées met en jeu l'appartenance à un groupe social, et lorsque l'abandon de ces idées comporte la déchéance du groupe lui-même.

Devons-nous considérer toutes ces réactions comme étant irrationnelles ? Tout dépend de la signification que nous donnons à ce terme. Les informations concernant des dangers et menaces sont certainement utiles dans une perspective de survie. Il n'est alors pas surprenant que nous soyons particulièrement enclins à écouter des informations portant ce type de contenu, qu'elles soient vraies ou fausses. L'adhésion à un groupe est une condition certainement propice à la survie pour une espèce sociale et culturelle comme la nôtre.

Notre équipement naturel – fruit de notre longue évolution – est formé par la pression sélective de collecter des informations utiles à notre survie plutôt que de se préoccuper de filtrer des connaissances scientifiquement prouvées[125]. Dans un sens, adaptatif, il est donc rationnel de donner la priorité à certains contenus, même s'ils peuvent se révéler faux à la lumière des faits rigoureusement établis.

3.2.4.5 Le rôle du prestige de la source

Si certains éléments du contenu des informations peuvent tromper nos défenses naturelles, ceci pourrait être vrai aussi par certaines caractéristiques des sources d'information. Un courant d'études s'intéresse aux mécanismes de l'apprentissage social, considéré comme très important pour expliquer l'accumulation de connaissances produites culturellement (Chudek & Heinrich 2011[126]). Plus généralement, comme la communication, l'apprentissage social est soumis à des règles. Ainsi, le choix de « maîtres » ou d'exemples à suivre et à imiter pour apprendre implique des mécanismes proches de ceux de la vigilance épistémique, qui déterminent l'évaluation de la qualité des informations. Robert Boyd, Peter Richerson, Joseph Heinrich et d'autres anthropologues cognitifs ont étudié les critères et les biais qui influencent l'apprentissage social, tels que l'âge, le sexe, l'état de santé, la langue, l'appartenance ethnique ou l'expertise du modèle (Boyd & Richerson 2005; Richerson & Boyd 2008; Chudek & Heinrich 2011[127]). Par biais, nous entendons ici un critère qui influence un choix, même de manière non consciente. Comme pour la vigilance épistémique, se baser sur des critères souvent indirects et à bas coût peut représenter une source potentielle d'erreur. Ceci est particulièrement évident dans le cas de l'expertise, difficile à évaluer directement et donc inférée à partir d'indices plus ou moins liés aux capacités réelles du modèle, telles que : les signes perceptibles de compétence, les signes de succès préalable, les signes de prestige, qui se traduisent par le fait que le modèle est suivi et copié par d'autres[128]. Ce dernier critère est considéré comme un indice particulièrement risqué. En vertu de ce biais, un individu peut en effet être suivi et imité – donc devenir influent – indépendamment de ses compétences réelles dans le domaine d'expertise recherché (Henrich & Gil-White 2001).

Chudek et al (2011) ont testé l'existence du biais de prestige chez l'enfant. Ils ont exposé des enfants d'environ trois ans à deux modèles. L'un est un modèle prestigieux selon la définition présentée plus haut : il a des *followers*, des assistants de l'expérimentateur qui le regardent attentivement. L'autre ne reçoit pas ce type d'attention de la part des assistants et est donc considéré comme non prestigieux. Les deux modèles, prestigieux et non prestigieux, performent différentes tâches : ils nomment des objets, expriment leurs préférences alimentaires pour différents aliments ou objets. Les résultats indiquent que le bais de prestige

est présent dans différents domaines d'apprentissage culturel (y compris les préférences alimentaires) mais sa force varie d'un domaine à l'autre[129]. Une revue récente de la littérature suggère cependant que l'influence du biais de prestige dans l'apprentissage social est réduite et modulée par différents facteurs tels que la difficulté de la tâche à apprendre, l'importance de la tâche et la présence d'autres biais de l'apprentissage social – notamment les biais liés aux contenus[130]. La manière dont les caractéristiques de la source d'information peuvent attaquer nos défenses naturelles est donc encore à explorer.

Nous savons que la communication à distance change la manière de déployer et donc de détecter les signes d'expertise. Certains signes visibles peuvent être ajoutés (*likes*, votes, *followers*) et éventuellement falsifiés. Il est donc crucial de comprendre comment ces signes sont produits et ce qu'ils traduisent réellement dans ces nouveaux environnements de communication et de partage de l'information.

3.2.5 L'incertitude comme moteur pour changer de position et suivre l'avis d'autrui

Plusieurs recherches montrent que la confiance dans ses connaissances (un aspect de la métacognition) est un modulateur fondamental de la capacité de changer de position et de l'adoption des positions d'autrui contre les siennes (Mercier 2017).

L'étude de Yaniv (2004) sur la disponibilité à changer de position et à prendre l'avis d'autrui, citée plus haut, montre par exemple que les participants prennent en compte leur degré de connaissance et utilisent cette auto-évaluation pour décider si doivent prendre ou rejeter l'avis d'autrui : plus ils pensent savoir, moins ils écoutent l'avis des autres ; moins ils s'estiment savants, plus ils sont prêts à modifier leur premier avis après avoir écouté celui des autres. Nous avons vu que, plus les avis des autres sont éloignés, moins ils influencent l'avis final, mais, encore une fois, cela dépend de la confiance que le sujet place dans ses propres connaissances[131].

L'influence de l'incertitude sur la susceptibilité d'accepter l'avis d'autrui a déjà été mise en évidence dans le comportement de bébés de dix-huit mois, devant décider de traverser ou non un passage difficile (Tamis-LeMonda et al. 2008). Lors d'une étude conduite sur environ vingt couples de sujets (mère et enfant), Tamis-LeMonda et al. (2008) observe que les enfants tendent à ignorer l'avis des mères qui les invitent à traverser une pente si la pente est perçue comme trop dangereuse ; ils ignorent également celles qui appellent à ne pas traverser si la pente est perçue comme facile. Dans la condition intermédiaire, une pente ni trop facile ni trop difficile, le comportement de l'enfant change selon l'encouragement maternel, avec plus de chances de traverser si la mère encourage que si elle dissuade[132]. Nous sommes donc en présence d'une sensibilité précoce à l'avis de l'autre en relation avec le degré de confiance que l'enfant a en lui-même et en sa perception du danger.
Le rôle de l'incertitude a aussi été exploré dans le cadre d'études sur l'apprentissage social et la transmission culturelle chez d'autres animaux[133]. Le problème consiste à déterminer quand un individu a intérêt à en imiter un autre (apprentissage social) plutôt que de chercher à apprendre par lui-même (apprentissage individuel). Comme toute autre forme de communication, l'apprentissage social est en effet avantageux seulement si certaines conditions sont données, telles que la fiabilité et l'expertise du modèle à imiter. Cela est vrai pour notre espèce comme pour les autres espèces sociales. À l'aide de modèles, les chercheurs ont montré qu'une condition fondamentale poussant à l'apprentissage social chez différentes espèces animales était l'incertitude. Ils ont appelé ce critère le « biais copie si incertain ». Ce biais a été mis en évidence non seulement à l'aide de modèles mathématiques mais aussi d'études empiriques sur des espèces non humaines (*Rattus norvegicus* : Galef, Whiskin & Dudley 2008 ; *Pungitus pungitus* : Van Bergen Coolen & Laland 2004[134]).

Cela donne une indication forte afin de chercher d'autres *building blocks* cognitifs de l'EC dans les capacités métacognitives liées à la perception de soi-même, soit comme compétent et

performant soit comme incertain, et à la perception ou à l'anticipation de l'erreur dans ses choix et décisions.

Avant d'analyser la littérature pertinente, terminons cette première partie en présentant les conséquences des thèmes traités pour l'éducation de l'EC.

3.2.6 La vigilance épistémique. Conséquences pour l'éducation

Les mécanismes appelé de « vigilance épistémique » nous permettent d'utiliser la communication avec les autres et de les transformer en sources d'information utiles, en filtrant l'information sur la base de la pertinence, de la plausibilité, de la vérité de ses contenus et de la fiabilité (bienveillance et expertise) des sources de l'information elle-même.

Ces mécanismes sont donc à la base de notre capacité d'évaluer la qualité épistémique de l'information afin de prendre une décision – une capacité que nous avons identifié avec l'EC. Pour cette raison, nous considérons les mécanismes de la vigilance épistémique comme des mécanismes à la base de l'EC. De plus, puisque ces mécanismes sont naturels, cela nous conforte dans l'idée selon laquelle l'EC est – aussi – naturel. Nous avons cependant mis en évidence les limites de la vigilance épistémique naïve et, par conséquent, de l'EC. Nous avons également souligné que ces mécanismes sont susceptibles de nous asmener en erreur en raison de la nature indirecte des critères utilisés pour évaluer contenus et sources d'information. Enfin, ces mécanismes ne sont pas sans faille et des brèches sont vite ouvertes dans nos défenses naturelles. Ceci est d'autant plus vrai que la pression sélective ayant permis le développement des mécanismes de la vigilance épistémique s'est probablement exercée dans des contextes assez différents de celui dans lequel nous vivons actuellement, notamment au niveau des systèmes d'échange et de partage de l'information.

L'éducation de l'EC a pour but de maximiser la capacité à évaluer la qualité de l'information en vue d'une prise de décision dans une variété de situations, actuelles et complexes.

Vue l'existence de mécanismes naturels d'évaluation des informations, développer l'EC signifie donc essentiellement l'outiller : exercer des critères plus sophistiqués que ceux utilisés de façon spontanée pour répondre à des situations plus complexes de façon correcte.
Il est donc nécessaire de prendre en compte les capacités naturelles et les critères spontanément utilisés pour l'évaluation des sources et contenus informationnels, et leurs limites.

Il ne s'agit pas de créer des capacités *ex novo*, ni de développer le raisonnement, la rationalité ou la pensée de façon générale. Il s'agit de façon beaucoup plus pratique de fournir des connaissances – sur ce qui constitue une source fiable dans un certain contexte, par exemple, ou sur ce qui constitue une preuve solide à l'appui d'une affirmation.

La notion d'EC adapté au contexte permet de distinguer un EC avancé d'un EC expert. L'EC avancé est un EC outillé pour répondre aux défis communs du quotidien. En théorie, tout le monde, dans des situations courantes, possède des connaissances propres à l'EC avancé. L'EC expert repose lui sur des connaissances spécialisées, qui ne sont pas utiles à tout le monde en toute circonstance, mais qui permettent d'évaluer correctement les informations dans des contextes spécifiques.

L'indication principale pour l'éducation consiste à fournir aux apprenants connaissances et critères utiles pour l'évaluation des informations :

- pour évaluer la pertinence d'une information portée à l'appui d'un argument ;
- pour évaluer la plausibilité des contenus d'information ;

- pour évaluer la validité de ces mêmes contenus de l'information, sur la base des preuves disponibles ;
- pour évaluer l'expertise de la source de l'information ;
- pour évaluer sa bienveillance (ses intérêts cachés et tout ce qui pourrait justifier ou faire penser qu'elle veuille nous tromper).

Ces outils acquis s'intègrent ensuite aux outils naturels, mais ils sont plus sophistiqués que ceux utilisés spontanément pour l'évaluation des contenus et des sources, et mieux adaptés au contexte présent.

L'évaluation de la pertinence et de la plausibilité d'une information dépend notamment du fait de posséder des connaissances dans le domaine concerné. Une idée apparaît plausible en raison de la base de connaissances dont nous disposons. Ainsi, la phrase « Ce crystal a le pouvoir de soigner grâce à son énergie » peut être jugée comme parfaitement plausible par quelqu'un qui ignore la notion scientifique d'énergie. L'idée que des pilules homéopathiques puissent soigner des maladies est plausible si nous ignorons leur réelle composition ou le fait que la mémoire de l'eau n'est pas une théorie validée scientifiquement.

De cette considération découle l'importance de ne pas disjoindre l'éducation de l'EC de l'acquisition de connaissances factuelles, c'est-à-dire des connaissances scientifiquement validées sur le monde physique, biologique, social, etc.

L'évaluation de la validité des contenus fait appel à la capacité à utiliser correctement les observations et faits ou preuves à l'appui. Il s'agit d'une tâche difficile, et c'est pour cette raison que, dans le cas des informations de seconde main, l'évaluation des sources prime souvent sur celle des contenus (sauf pour l'évaluation de la pertinence et de la plausibilité, c'est-à-dire de la cohérence avec ses propres connaissances). Imaginons que nous voulions évaluer une affirmation telle que « Ce vaccin homéopathique permet de prévenir la grippe » sur la base des preuves à l'appui. Pour cela, nous devrions savoir que les observations individuelles (anecdotes) ont une valeur de preuve très faible. De plus, nous aurions besoin de connaître la valeur ajoutée, pour éviter les biais d'observation et les facteurs confondants, de tests expérimentaux avec des groupes de contrôle et une randomisation, et de la répétition de ces mêmes tests. Grâce à ces connaissances, il serait possible de comparer la valeur de preuve d'un témoignage fourni par un proche à celle de preuve d'études scientifiques ayant recours aux méthodes citées. C'est justement ce type de connaissance qui peut enrichir notre EC naturel et nous rendre plus aptes à distinguer des contenus informationnels sur la base de la qualité des preuves à l'appui.

Ces connaissances permettent en outre d'obtenir un objectif supplémentaire. Devant la difficulté de pratiquer ce type d'évaluation des contenus sur une base quotidienne, comprendre quelles méthodes permettent d'obtenir les preuves les plus certaines permet d'identifier des classes de sources d'information plus fiables sur la base de leur expertise. Les scientifiques, qui mettent en pratique des méthodes rigoureuses pour établir des relations de cause à effet, méritent notre confiance en vertu de cette même pratique. Connaître les méthodes de validation des connaissances (de l'observation rigoureuse à l'expérimentation, au *peer reviewing* et aux autres stratégies de contrôle social sur les résultats individuels) permet ainsi de mieux identifier non seulement les contenus plus solides, mais aussi les sources plus fiables.

De cette considération découle la nécessité d'exploiter les enseignements scientifiques afin de fournir une connaissance de plus en plus approfondie concernant les stratégies et méthodes permettant de mettre sous contrôle les principaux biais d'observation et d'articuler correctement théories et faits observés. Il n'est pas nécessaire – pour faire preuve d'EC – d'être capable de mettre en place une expérience ou de conduire une observation rigoureuse à la manière d'un scientifique. Connaître ces stratégies et méthodes, leurs raisons et leurs

conséquences permet cependant distinguer des contenus d'information appuyés sur des preuves de contenus non validés, et des sources fiables en vertu des méthodes utilisées pour aboutir à une conclusion de sources moins fiables en matière d'expertise.

Une certaine littérature scientifique, relative aux méthodes et aux rouages de la science, permet entre autres de reconnaître la spécificité de la connaissance scientifique et d'identifier la science comme une source privilégiée de connaissances concernant les phénomènes naturels.

L'évaluation de la validité des contenus fait appel à la capacité d'utiliser correctement les observations et faits ou preuves à l'appui.

3.3 Les bases cognitives de l'EC : la sensibilité métacognitive, la confiance

L'évaluation de l'information pour faciliter une prise de décision ne se limite pas à l'évaluation de l'information de seconde main.

Reprenons l'exemple du restaurant. Cette fois, au lieu de demander l'avis de quelqu'un, nous choisissons de nous baser sur nos observations pour choisir « le meilleur restaurant ». Nous observons la queue devant le restaurant, son apparence, l'aspect des plats. Sur la base de ces informations récoltées de première main, nous prenons notre décision. Cet exemple permet de mettre en parallèle l'évaluation des informations de seconde main avec celle des informations directes, de première main.

Dans les deux cas émerge en effet la question de la confiance. Dans le cas de l'information de seconde main, il s'agit de la confiance (*trust*, en anglais) que nous avons dans nos sources ou dans le contenu de l'information qui nous est délivrée, en vertu de l'évaluation par critères que nous venons de décrire. Dans le cas de l'information de première main, que nous collectons à l'aide de nos sens ou qui est stockée dans notre mémoire, il s'agit de la confiance que nous avons dans nos propres idées, dans nos perceptions, nos souvenirs et dans nous-mêmes en tant que sources : « Nous sentons-nous sûrs de notre opinion, de notre décision ? Avons-nous un doute, ressentons-nous une incertitude ? »

Il suffit de se trouver dans un endroit sombre et inconnu pour expérimenter une baisse immédiate de confiance dans les informations qui nous sont fournies par nos organes sensoriels. Dans ces cas, nos pas se font plus incertains, nous ralentissons notre marche pour accumuler plus d'informations au fur et à mesure et ainsi éliminer l'hypothèse de possibles obstacles interposés. Nous savons qu'il est plus facile de se tromper dans ces conditions. Nous n'avons même pas besoin de nous consulter avec nous-mêmes pour savoir que notre ressenti en termes d'incertitude est affecté par la situation. Même sans que cela devienne explicite, nos décisions sont affectées par nos sentiments de confiance par rapport aux informations disponibles et à nos chances de nous en sortir sans casse.
Cet exemple de la marche dans le noir est un exemple parmi d'autres où évaluation de l'information, décision et confiance se trouvent liées de façon silencieuse mais sûre.
Comment savons-nous quand il est raisonnable de nous faire confiance ? Comment réussissons-nous à estimer nos chances de nous tromper, nos possibilités de commettre une erreur plutôt que de prendre la bonne décision ? Cette estimation est-elle plus correcte ? Notre forte confiance correspond-elle au fait de ne pas nous tromper alors qu'une faible confiance correspondrait plutôt à des situations où, en effet, nous nous trompons plus facilement ?

Précisons que, dans les pages suivantes, nous entendons par là la confiance en nos propres perceptions, connaissances, opinions et croyances, décisions et non une confiance en soi plus

générale. Nous étudions aussi comment la confiance dans ces états mentaux influence la décision et l'apprentissage. Par décision, nous entendons un état mental consistant à opter pour un choix alors que plusieurs options sont possibles. Une décision peut donc être de nature purement perceptive : confrontés à un stimulus sonore ambigu, nous décidons qu'il s'agit d'un bruit de pluie ; face à un stimulus visuel incertain, perçu dans un environnement sombre, nous décidons qu'il s'agit d'un chat. L'apprentissage est ici défini comme la mise à jour des prédictions, représentations, décisions préalables à la lumière de nouvelles preuves (évidence, retour d'expérience, nouvelles observations, nouveaux stimuli). Il est donc apparenté aux exemples de changement de position et d'opinion (en situation d'incertitude et/ou à la lumière de preuves provenant de l'extérieur) discutés dans la section précédente.

La littérature sur la confiance est, depuis une dizaine d'années, en rapide expansion. Cette littérature mobilise un large éventail de disciplines de la famille des sciences cognitives : philosophie, neurosciences – neuro-imagerie, modélisation –, psychologie du développement, psychologie expérimentale, psychologie du raisonnement et sociale (Proust 2013 ; Fleming & Lau 2014 ; Grimaldi, Lau & Basso. 2015 ; Goupil & Kouider 2019). Elle se situe dans un cadre plus large : celui des études concernant la métacognition et, en particulier, la métacognition implicite et la sensibilité métacognitive.

3.3.1 Aspects de base de la métacognition

Pouvons-nous faire confiance à cette information ? Est-ce raisonnable d'utiliser cette information pour prendre une décision ? La décision que nous venons de prendre est-elle justifiée par des informations disponibles ? Posées de manière explicite, à nous-mêmes, ces questions renvoient à un groupe d'états mentaux particulier gérés par des processus métacognitifs. Une définition classique de métacognition est celle de « *thoughts about one's own thoughts and cognition* » (Flavell 1979). Ainsi définie, la métacognition apparaît comme une forme de cognition de second ordre (des états cognitifs ayant pour objet d'autres états cognitifs du même individu) (Grimaldi, Lau & Basso 2015[135]). Le rôle fonctionnel de la métacognition serait double : celui de surveillance des états mentaux (Dunlosky & Metcalfe 2009) et celui de contrôle sur ces mêmes états. La métacognition est donc définie comme la capacité de monitorer et de contrôler la cognition, ou la capacité de réfléchir sur nos représentations mentales afin de réguler la cognition et optimiser l'apprentissage. Elle inclut plusieurs processus comme la régulation de l'attention, le choix de se former ou de s'informer, l'action de repérer et de corriger ses erreurs (Proust 2019[136]).

Les domaines d'application de la métacognition sont vastes et variés et, par conséquent, les acceptions du terme peuvent légèrement changer d'un auteur à un autre (Fleming, Dolan, Frith 2012[137]). Même avant que la métacognition ne devienne un thème de recherche en soi en psychologie, notamment grâce à Flavell, des études sur la mémoire ont exploré des expériences métacognitives comme la sensation d'avoir un mot sur le but de la langue (ToT : *tip of the tongue*) et la sensation de savoir (FOK : *feeling of knowing*). La métamémoire est encore aujourd'hui un champ de recherche assez suivi. Cependant, les études se sont entretemps étendues au-delà des frontières de la psychologie. La métacognition est aujourd'hui un objet de recherche multidisciplinaire, particulièrement représenté dans les domaines de la psychologie éducative et de la psychologie cognitive, mais aussi en psychologie clinique, animale et en neurosciences (Norman et al. 2019). Parmi les fonctions les plus explorées en relation avec la métacognition se trouvent l'apprentissage (par exemple, Bransford, Brown, Cocking 2000), la relation de la connaissance à la croyance (par exemple, Wellman, Cross & Watson 2001) et la décision comme, par exemple, la décision perceptive : « Est-ce que l'objet que j'entrevois dans le noir est bien un chat ? » « Quel objet produit ce bruit ? » « J'ai l'impression d'entendre la pluie tomber mais en réalité il ne pleut pas. » (Norman et al. 2019). Dans des situations de prise de décision, la métacognition intervient notamment à deux niveaux : permettre d'anticiper les chances d'erreur et jauger la fiabilité (le degré de certitude) d'une décision/information par rapport à d'autres. En vertu de ces

composantes, la métacognition permet de guider notre comportement dans des situations complexes : par exemple, en présence de multiples sources d'information ou d'informations ambiguës qui se prêtent à différentes interprétations. Ainsi, estimer correctement la confiance que mérite une certaine information permet de décider s'il est nécessaire de rechercher de nouvelles informations ou si les informations disponibles suffisent. Nous parlons alors de « confiance métacognitive » et la définissons comme un jugement de probabilité ayant pour objet une action ou la probabilité qu'un jugement donné soit correct. La confiance métacognitive pourrait également revêtir un rôle fonctionnel au niveau de la cognition sociale, par exemple dans le cas de la prise de décision collective, afin de résoudre les éventuels conflits d'opinion en faveur de celles plus fiables, ou de permettre de changer d'opinion en faveur d'une source plus fiable (Fleming, Dolan & Frith 2012[138]). (Dans la partie précédente, nous avons anticipé, en vertu des connaissances issues d'études comportementales, que l'incertitude était un moteur de changement de position, c'est-à-dire de mise à jour de ses opinions et jugements pour adopter ceux d'autrui.)

3.3.2 La métacognition implicite

La métacognition apparaît dans la définition de Flavell comme un état intentionnel dans lequel nous prenons en compte, volontairement, nos états mentaux de façon réflexive.

Cependant, le concept de métacognition est plus large que celui de métacognition explicite ou déclarative (avec transmission verbale de croyances, de théorie naïves ou expertes sur la cognition) ou conceptuelle. Classiquement en effet, nous distinguons les connaissances métacognitives (je sais dans quel état cognitif je me trouve, je comprends mes capacités cognitives, je sais ce qui me permet de les améliorer), les stratégies métacognitives (je mets en place des actions qui vont me permettre d'améliorer mes performances, ou en tout cas de les contrôler), les expériences ou sentiments métacognitifs (j'éprouve la sensation d'être près du but, ou en difficulté dans une tâche, j'ai un sentiment de familiarité avec un exercice, le sentiment de pouvoir réussir) (Flavell 1979 ; Efklides 2008, 2011). Les sentiments métacognitifs produits en cours d'activité ont un rôle important dans la prédiction d'échec ou de succès d'une tâche, donc dans l'auto-évaluation métacognitive qui précède la réponse à la tâche elle-même (Proust 2019, Koriat 1993[139]). Ces sentiments émergent comme les pointes visibles de l'iceberg à partir de processus ayant lieu silencieusement dans notre cerveau.

Le cerveau a en effet recours à des indices prédictifs qui lui permettent d'estimer ses chances de réussir dans une tâche – d'estimer sa confiance dans sa capacité à résoudre un problème, à prendre la bonne décision, ses chances de répondre correctement ou de commettre une erreur. Ces indices sont indirects : ce sont des « heuristiques prédictives ». Il peut s'agir de la rapidité avec laquelle l'information liée à la tâche est traitée (heuristique de fluence) ou de la cohérence entre la réponse produite et d'autres représentations activées par la trace (heuristique de cohérence). Au niveau neural, nous identifions comme mécanisme de la confiance la convergence entre différents groupes de neurones vers une seule et unique décision (ampleur de l'activation produite par une tâche) (Proust 2019[140]). Les sentiments métacognitifs appartiennent donc à une forme de métacognition procédurale, implicite, sub-personnelle, silencieuse – qui n'a pas besoin de mots pour s'exprimer et agir sur le comportement.

3.3.3 Trajectoires développementales de la métacognition

La métacognition explicite, verbalisée, mature progressivement au cours de l'enfance. Par exemple, l'enfant sait exprimer correctement ce qu'il sait ou ne sait pas autour de sept-huit ans. Sa capacité à estimer la difficulté d'une tâche cognitive s'améliore au cours des premières années d'école. La sensation de savoir et la capacité de juger son apprentissage s'améliorent aussi, mais de manière moins prononcée. Le développement de stratégies

appropriées à la difficulté de la tâche occupe la dernière partie du parcours de l'école primaire.

Néanmoins, même des enfants beaucoup plus jeunes se montrent capables d'estimer leur niveau de connaissance – implicitement – et de chercher de l'information supplémentaire – ou de s'abstenir de donner une réponse en cas d'incertitude (voir plus loin). Des estimations de certitude et d'incertitude sont donc présentes sous une forme ou une autre indépendamment du fait de pouvoir être exprimées verbalement. Nous retrouvons des signes de métacognition implicite chez l'adulte, aussi bien que chez le jeune enfant d'âge préverbal et chez l'animal non humain (Palmer, David & Fleming 2014 ; Smith, Shields & Washburn 2003 ; Grimaldi, Lau & Basso 2015[141]).
Leur expression est comportementale et mise en évidence pas des dispositifs expérimentaux silencieux tels que :

- la possibilité donnée au sujet de rechercher d'autres informations en plus de celles en sa possession ;
- la possibilité de ne pas faire un choix entre deux options, mais de sortir de l'expérience (*opt-out*) de manière à minimiser les pertes tout en renonçant à la possibilité de gagner une plus grande récompense : le sujet, de façon prospective, décide à l'avance de ne pas donner un certain type de réponse s'il sait que la récompense est plutôt associée à une réponse correcte (grande récompense) et à pas de réponse (petite récompense), mais qu'il perd la chance d'une récompense s'il donne une réponse incorrecte.

L'idée étant que le sujet estime son niveau de connaissance avant de répondre et agit en conséquence pour maximiser les bénéfices) (Norman et al. 2019 ; Shea et al. 2014). Ces dispositifs permettent de dépasser le préjugé selon lequel la métacognition est limitée aux organismes qui maîtrisent le langage.

Leur nature entraîne cependant plusieurs interprétations des résultats obtenus – notamment lorsque les tests sont dispensés sur des animaux – et maintient ouvert le débat sur la relation de la métacognition à la conscience. Les études sur la confiance chez l'animal sont par exemple souvent critiquées selon des bases méthodologiques. En effet, il est difficile d'exclure que d'autres mécanismes que l'évaluation de la confiance sont en jeu. Il est cependant accepté de façon assez consensuelle que les êtres non humains (animaux) sont capables d'estimer leur confiance en une certaine décision ou connaissance et, plus généralement, qu'ils possèdent des mécanismes métacognitifs de type implicite (Fleming, Dolan & Frith 2012[142], Shea et al. 2014[143]).

Certains auteurs ont proposé de distinguer entre, d'un côté, processus métacognitifs de type automatique, implicites et basés sur des indices indirects et, de l'autre, processus métacognitifs plus exigeants, supra-personnels (Shea et al. 2014). Les premiers permettent de contrôler l'activité neurale même à des niveaux très bas du traitement, en prenant en compte l'incertitude relative à chaque représentation, et de contrôler l'activité de différents systèmes perceptifs et moteurs. Les seconds permettent d'exprimer cette confiance de manière explicite, voire verbale. Ils ont donc un rôle fonctionnel particulièrement important dans la coopération entre individus, lorsqu'il s'agit de prendre des décisions à plusieurs en considérant l'incertitude et la fiabilité de chacun. D'autres auteurs – nous le verrons plus loin – ont proposé d'introduire des processus neuraux qui lisent l'information concernant la confiance à partir des propriétés de l'activité des neurones impliqués dans une décision (Meyniel et al. 2015). Dans les deux cas, ce qu'il est important de retenir concerne l'omniprésence des processus métacognitifs implicites et leur dépendance à des indices indirects et l'idée selon laquelle d'autres processus peuvent entrer en jeu en exploitant les processus implicites et permettant *in fine* de les améliorer.

3.3.3.1 Les mécanismes précoces de la sensibilité métacognitive

Goupil & Kouider (2019) proposent une synthèse des connaissances actuelles concernant la métacognition implicite chez les jeunes enfants. Cette synthèse montre que des enfants de trois à cinq ans sont capables par des moyens non verbaux d'exprimer une confiance bien calibrée par rapport à leurs décisions, et que des enfants de douze à vingt mois manifestent des capacités métacognitives implicites. Ces résultats contredisent une vision bien affirmée du développement de la métacognition selon laquelle cette fonction est rare pendant l'enfance (Flavell 1999, Schneider 2008, Sodian et al. 2012).

Les enfants de trois à cinq ans peuvent en effet être entraînés à utiliser une échelle de confiance non verbale (images d'un enfant confiant ou qui doute). À l'aide de ce type de dispositif, ils peuvent ensuite exprimer la confiance qu'ils portent aux informations en leur possession (confiance rétrospective, explicite). Nous constatons que les réponses correctes sont corrélées avec une plus grande confiance et celles non correctes avec une moindre confiance : les enfants de cet âge montrent donc qu'ils sont capables d'évaluer correctement leurs performances et d'avoir une sensibilité métacognitive.

La notion de sensibilité métacognitive est une notion normative : un sujet fait preuve de sensibilité métacognitive quand il est confiant dans la décision qu'il a prise ou à prendre et que cette confiance s'associe à des performances effectivement correctes. De manière similaire, sa confiance métacognitive est bien calibrée quand sa confiance est basse et que les performances sont insatisfaisantes. À l'opposé, dans les cas de confiance mal calibrée, confiance et réalité divergent (Fleming & Lau 2014[144]).

Une distinction classique au sein de la confiance métacognitive est celle entre métacognition rétrospective et métacognition prospective (Grimaldi, Lau & Basso 2015[145]). La première est mesurée de la manière suivante : on demande aux sujets qui participent à l'expérience d'évaluer leur succès à une tâche i que cette dernière ait été effectuée[146]. La seconde est mesurée de la manière suivante : on propose une tâche et on demande aux sujets d'anticiper leurs chances de succès ; dans ce cas, les suejts peuvent décider si répondre ou non (comme dans le cas du paradigme *opt-out* décrit plus haut)[147]. Les paradigmes de type prospectif sont donc plus souvent utilisés pour étudier la métacognition implicite.

Des paradigmes de type *opt-out* ont permis de montrer que la capacité d'estimer la probabilité de commettre une erreur est déjà présente chez l'enfant de vingt mois et que la sensibilité métacognitive par rapport à la justesse de ses décisions est présente dès douze mois. Voici quelques exemples.

Des enfants de vingt mois doivent se souvenir de l'endroit où un jouet a été caché. Pour varier la difficulté, un temps plus ou moins long est donné entre le moment où le jouet est caché et celui où l'enfant doit le retrouver. Les enfants du groupe témoin sont forcés au choix alors que les enfants du groupe expérimental ont la possibilité de faire un *opt-out* en demandant l'aide d'un adulte. Le taux de réponses correctes par rapport au recours à l'aide de l'adulte indique que les enfants estiment correctement le risque de faire une erreur, donc leur incertitude (Goupil, Romand-Monnier Kouider 2016[148]).
Pour calibrer la confiance dans la décision prise, Goupil et Kouider (2016) ont utilisé le même paradigme mais avec une variation : mesurer la persistance dans la recherche d'un objet caché dans une boîte selon le degré de certitude que l'objet se trouve réellement dans la boîte. Les auteurs ont mené leur étude avec des enfants de dix-huit mois ; ces derniers observent un adulte cacher un objet dans une boîte. Le délai entre le moment où l'objet est caché et celui où l'enfant doit pointer avec son doigt dans quelle boîte l'objet est caché varie. Selon les auteurs, l'incertitude grandit avec l'augmentation du délai car les performances (retrouver l'objet caché dans une boîte plutôt que dans une autre) baissent avec le temps écoulé. Ils ont alors proposé l'hypothèse selon laquelle la persistance dans la recherche de l'objet caché dans la boîte est d'autant plus forte quand l'enfant est sûr de trouver l'objet dans la boîte qu'il a

indiquée, donc qu'il est confiant dans son choix. Ils ont constaté que, en effet, les enfants cherchent plus longtemps – persistent plus – dans le cas où leur réponse est correcte. Ils ont interpreté cela comme un signe de correct calibrage de la confiance métacognitive dans la décision (Goupil & Kouider 2016[149]). Un autre test de persistance consiste à cacher l'objet désirable (un biscuit) soit dans une boîte facile à ouvrir soit dans une boite comportant un couvercle scellé, qui nécessite l'aide d'un adulte pour être ouverte. De la même manière, le délai temporel entre le moment où l'objet est caché et celui où il faut le retrouver varie (trois à douze secondes). Si l'objet est caché dans la boîte scellée, l'enfant – convaincu que l'objet se trouve à l'intérieur – doit demander l'aide de l'adulte pour l'ouvrir. S'il n'est pas vraiment convaincu, il essaye d'ouvrir l'autre boîte pour voir. La persistance dans son choix dépend donc à la fois du fait d'avoir fait le choix correct et du temps écoulé (Goupil & Kouider 2016[150]).

Le même type de capacité a été mesurée chez des enfants de cinq à huit ans en utilisant le paradigme dit de « pari après décision ». Les enfants observent des images contenant des objets (points) en quantité variable. Suite à l'estimation (« Laquelle des deux images contient le plus d'objets ? »), ils doivent exprimer leur confiance dans leur jugement (« Combien es-tu sûr de ta réponse ? »). L'expression de confiance passe par le choix entre un émoticône représentant un visage souriant et un autre représentant un visage incertain. Les résultats montrent une corrélation positive entre difficulté de la tâche et performances, et entre ces dernières et expression de confiance (sensibilité métacognitive). Dans le cadre de ce test, la sensibilité métacognitive s'améliore avec l'âge, de cinq à huit ans (Vo et al. 2014).

Enfin, d'autres preuves confirmant la présence d'une forme précoce de sensibilité métacognitive chez l'enfant sont constituées par la neuro-imagerie chez l'adulte et dans d'autres espèces, comme les rats et les singes (Goupil & Kouider 2019). Chez l'adulte, la métacognition implique l'activation du cortex préfrontal : la confiance en la décision semble être encodée au niveau du cortex préfrontal ventromédian ; le monitoring de l'erreur (décalage entre le choix erroné venant d'être effectué et le choix qui aurait dû être fait sur la base de l'évidence disponible) est représenté au niveau du cortex cingulaire antérieur et par un signal caractéristique en EEG (*Error-Related Negativity*) (Goupil & Kouider 2016). Ces structures sont déjà actives pendant l'enfance (par exemple, la composante ERN est présente en EEG après un choix incorrect chez l'enfant de douze mois), ce qui confirme que la sensibilité métacognitive n'est pas limitée à des capacités explicites chez l'adulte.

Les auteurs suggèrent donc l'existence d'un système métacognitif précoce, ou « core métacognition », qui serait présent chez l'enfant avant qu'il soit capable d'en réaliser un rapport verbal et également chez d'autres espèces. La fonction de ce système précoce serait celle d'évaluer et de réguler la cognition de façon automatique et implicite, notamment d'évaluer la qualité d'une représentation : la probabilité (*likelihood*) que la représentation soit correcte, en raison des preuves disponible (de type perceptif). Sur la base du fait que le cortex préfrontal mature plus lentement que d'autres régions cérébrales comme celles impliquées dans la décision perceptive, nous pouvons supposer que la maturation de la core métacognition est elle aussi lente. Le *core metacognition system* présenterait aussi d'autres limites propres : le fait que l'évaluation passe par l'utilisation d'indices, comme le temps pris pour la décision, plutôt que de calculs plus fiables. Le fait que les évaluations implicites ne peuvent pas être transmises verbalement rend impossible d'exploiter la core métacognition dans un échange social, pour communiquer son niveau de certitude. Pour en arriver là, la core métacognition a besoin de la maturation et de l'entrée en jeu d'autres capacités, comme la capacité de porter l'attention sur les représentations métacognitives, ce qui les rend explicites ou conscientes. Nous sommes donc en présence d'un problème d'attention et de saillance de la représentation par rapport à des stimuli externes. Nous pouvons aussi supposer que l'attitude surconfiante des enfants (et des adultes) qui répondent systématiquement « oui je sais » est liée à l'incapacité à sélectionner (attention sélective) la bonne source d'information : le signal d'erreur plutôt que le désir de réussir. Il s'agirait donc d'un défaut de capacités exécutives. Les auteurs questionnent le rôle de la culture et des interactions sociales sur le

développement des systèmes métacognitifs (apprentissage d'expressions pour exprimer ses états métacognitifs, par exemple).

3.3.4 Confiance et cerveau

La notion de confiance a récemment passé la barrière de la psychologie pour intéresser les neurosciences. Cet intérêt se traduit dans un effort d'identification des régions cérébrales et, plus généralement, des substrats neuraux, qui corrèlent avec le comportement observable de confiance métacognitive. Cet effort passe par des études en neuropsychologie : la TMS, la MRI structurelle, la fMRI, l'inactivation pharmacologique sélective. Il n'existe pas de consensus pour le moment sur la manière dont la confiance serait encodée au niveau neural : soit *via* des modules dédiés et séparés de ceux de la prise de décision (possiblement avec des représentations cérébrales multiples et en partie différentes en raison du contenu), soit plutôt *via* les circuits dédiés à la prise de décision (Grimaldi, Lau & Basso 2015[151]). L'hypothèse des réseaux spécialisés semble dominante dans la littérature, notamment pour sa capacité à expliquer les cas de disjonction entre performance (décision) et estimation de la confiance (Grimaldi, Lau & Basso 2015 ; Meyniel, Signamn & Mainen. 2015).

Les neurosciences se sont également penchées sur la recherche d'une définition de la confiance en termes purement neuraux et computationnels. Celle-ci passe notamment par la notion de cerveau bayésien et par des modèles du cerveau comme organe essentiellement probabiliste (Kepecs & Mainen 2012 ; Meyniel, Sigman & Mainen 2015 ; Grimaldi, Lau & Basso 2015).

3.3.4.1 Confiance et cerveau bayésien

L'hypothèse du cerveau bayésien est aujourd'hui tenue en grande considération afin d'expliquer comment le cerveau représente la réalité à partir de stimuli relativement pauvres, incertains ou ambigus (Pouget et al. 2013[152]). Ou encore, comment il prend des décisions (y compris perceptives) et comment il les met à jour à partir de connaissances préalables (apprentissages). Les modèles bayésiens permettent de rendre compte d'une variété de phénomènes observés, comme les illusions perceptives et les biais mais aussi les modalités d'apprentissage statistique du bébé (Meyniel 2016). Ces modèles s'inspirent de la règle de Bayes : un outil mathématique introduit par le révérend Bayes permettant d'exprimer la probabilité qu'une certaine hypothèse soit vraie selon les données disponibles, la plausibilité *a priori* de l'hypothèse elle-même (par rapport à d'autres connaissances) et la probabilité d'observer les données présentes si l'hypothèse est vraie[153]. Selon l'hypothèse du cerveau bayésien, chaque représentation que le cerveau se fait de l'environnement est donc en réalité exprimée en termes probabilistes et qu'elle prend en compte non seulement le stimulus présent mais aussi des *a priori* qui influencent la manière d'interpréter le stimulus.

Voyons maintenant comment ce modèle permet d'intégrer la notion de confiance au sein même du fonctionnement cérébral[154].

Prenons l'exemple d'une décision perceptive où le cerveau se trouve face à une ligne inclinée : de combien la ligne est-elle inclinée ? L'aspect fondamental du cerveau qu'il est important de prendre en considération ici est qu'une population de neurones – et non un seul neurone – répond au stimulus (la ligne inclinée). Chaque neurone qui s'active en réponse à la ligne le fait cependant d'une manière qui lui est propre. Ainsi, chaque neurone répond de manière privilégiée à une certaine inclinaison de la ligne : par exemple, il réagit avec une activation plus forte si la ligne a une inclinaison de 45°. D'autres neurones répondent aux lignes orientées à 45° mais de façon plus faible.

D'autres encore ne répondent pas aux lignes orientées à 45°, mais seulement à d'autres orientations. La distribution globale des activations dans la population de neurones correspond alors à une estimation de la probabilité que le stimulus présent ait une certaine orientation – par exemple 45° d'inclinaison (Knill & Pouget 2004 ; Meyniel, Sigman & Mainen 2015). La décision perceptive (la ligne est orientée à 45°) est une décision probabiliste. Si nous regardons la variance qui existe à l'intérieur de cette distribution, nous obtenons également une mesure implicite de l'incertitude qui accompagne l'estimation de cette probabilité : plus la variance est élevée, plus il existe de l'incertitude. L'incertitude est ici considérée comme l'envers de la confiance, donc nous pouvons dire que dans la réponse cérébrale au stimulus se trouve aussi inscrite une évaluation implicite de l'incertitude et de la confiance (Meyniel, Sigman & Mainen 2015). Pouget, Drugowitsch & Kepecs (2016[155]) définissent plutôt la confiance comme l'estimation de la probabilité que le choix perceptif effectué soit correct, sur la base des preuves (stimuli) disponibles, mais la distinguent des notions de certitude et d'incertitude, qui n'impliquent pas nécessairement un choix. Cependant, cette différence de définition n'est pas significative ici. Il existe en effet d'autres modèles, toujours de nature bayésienne, permettant de représenter une décision perceptive au niveau cérébral et de définir la confiance au niveau neural. Nous nous limitons au modèle discuté, car le but n'est pas de fournir une image complète des approches neurales de la confiance (pour cela, nous pourrons consulter : Dunlosky & Metcalfe 2008 ; Kepecs & Mainen 2012 ; Middlebrooks & Sommer 2012 ; Shadlen & Kiani 2013 ; Fleming & Frith 2014 ; Fleming & Lau 2014 ; Grimaldi, Lau & Basso 2015 ; Meyniel, Sigman & Mainen 2015 ; Pouget et al. 2016), mais de présenter le principe général selon lequel il est possible de représenter la confiance au niveau neural – subpersonnel – comme une conséquence de l'état d'activation de population de neurones, dans le cadre de modèles probabilistes du fonctionnement cérébral. Un concept psychologique, celui de confiance, se trouve donc traduit en termes purement neuraux, et purement implicites (voir aussi, à ce propos, Ott, Masset & Kepecs 2018[156]).

Meyniel et al. (2015) parlent de confiance distributionnelle pour présenter cette manière de concevoir la confiance au niveau neural. La confiance distributionnelle est indissociable de la décision. Dans notre exemple, l'incertitude liée à l'interprétation de l'orientation de la ligne est indissociable avec la probabilité que la ligne ait une certaine inclinaison. Une interrogation concerne la possibilité que ce type de description s'applique au-delà du cas spécifique de la décision perceptive à des jugements de confiance métacognitive se situant à un plus haut niveau computationnel, comme celles décrites dans la littérature en psychologie (Meyniel, Sigman & Mainen 2015[157] ; Ott, Masset & Kepecs 2018[158]).
Une autre interrogation, que nous avons anticipée plus haut en évoquant les circuits spécialisés pour la confiance métacognitive, concerne la possibilité que d'autres processus et mécanismes existent et effectuent une lecture de certaines propriétés de la distribution de l'activation (Meyniel, Sigman & Mainen 2015). Ces processus et mécanismes sont réunis sous le terme de « *read out* » ou « extraction de la confiance ». Par exemple, nous pouvons faire l'hypothèse de l'existence de mécanismes reliant séparément la direction vers laquelle pointe la distribution des activations et la variance dans cette même distribution (ou d'autres mécanismes).

Postuler l'existence d'un processus de *read out* permet d'obtenir – à partir d'une distribution similaire d'activations au sein d'une même population de neurones – une représentation de la confiance séparée et spécifique par rapport à la représentation de la décision (« confiance sommative » vs. « confiance distributionnelle », selon la terminologie utilisée par Meyniel, Sigman & Mainen 2015). Les deux processus sont à la fois unis (parce qu'ils ont lieu dans une même population de neurones) et disjoints (parce qu'ils sont lus par deux mécanismes différents[159]).

L'hypothèse du *read out* permet ainsi d'expliquer certains résultats obtenus en neurosciences selon lesquels, au niveau des activations cérébrales, les circuits de la prise de décision et ceux

de la confiance sont en partie superposables, mais non totalement (Meyniel, Sigman & Mainen 2015[160]).

Selon Meyniel, Sigman & Mainen (2015), le *read out* est le processus en vertu duquel il devient possible de fournir une estimation unique, scalaire, de la confiance que nous avons dans une décision, perception ou opinion. Quand, lors d'un test verbal ou explicite, un sujet exprime son niveau de confiance, il fait référence au résultat du *read out* de confiance qui accompagne celui de la décision. Quand il décide de fournir ou de ne pas fournir une réponse dans un paradigme de test de type *opt-out* ou pari, il fait référence de façon implicite au même *read out* unique.

Il n'y a pas d'obligation de *read out* une fois que des informations sont codées de manière probabiliste. La confiance peut rester encapsulée au sein du circuit qui traite l'information en question. Mais, si ce *read out* est effectué, il devient possible d'exprimer et donc de partager et de comparer la confiance sommative. Ce partage peut se faire à l'intérieur d'une même architecture cognitive – d'un cerveau, mais aussi entre plusieurs régions qui codent pour différents types de stimuli ou modalités sensorielles et autres. Le partage devient aussi possible entre individus. Les processus de *read out* pourraient donc jouer un rôle fonctionnel dans les situations coopératives, pour atteindre une décision collective optimale prenant en compte les incertitudes de chacun et, plus généralement, dans toutes les situations complexes où de multiples sources d'information et connaissances sont à prendre en compte (Meyniel, Sigman & Mainen 2015).

3.3.4.2 Confiance, sensibilité à l'erreur et cerveau prédictif

Le cerveau apparaît dans plusieurs modèles neurocomputationnels comme un organe non seulement probabiliste mais prédictif (Rao & Ballard 1999). Le modèle prédictif le plus simple représente le processus de détection d'un signal ou d'apprentissage de façon hiérarchique et linéaire : face à un *input* (stimulus), le cerveau émet une prédiction ; celle-ci est ensuite comparée avec la réalité et produit un signal d'erreur (erreur de prédiction). Ce signal déclenche une mise à jour de la prédiction initiale (apprentissage).

Cette correction peut être d'autant plus grande que l'erreur de prédiction est importante. À la base de tout apprentissage ou traitement du signal, il y aurait donc une prédiction, croyance ou représentation initiale, puis un signal d'erreur si la prédiction est incorrecte (voir : Rescorla & Wagner 1972 ; Sutton & Barto 1998). Dans les années 1990, le rôle de la prédiction (et de la violation d'une prédiction) dans l'apprentissage est de plus appuyé par des données et modèles en neurosciences. Par exemple, Schultz, Dayan & Montague (1997) identifient des groupes de neurones qui, dans le cas de l'apprentissage par récompense et punition, signalent les erreurs de prédiction (neurones dopaminergiques dans VTA et *substantia nigra*[161]) : lorsque l'animal reçoit une récompense, son système dopaminergique s'active ; lorsqu'il comprend une association entre stimulus et récompense, cet apprentissage modifie l'activation du système dopaminergique, qui s'active ensuite avant même l'arrivée de la récompense, et signale ainsi une attente de récompense. Suite à l'arrivée d'un stimulus, l'organisme réalise donc une prédiction selon laquelle la récompense devrait bientôt arriver, et cette prédiction est associée à une certaine activation neurale avec décharge de dopamine. Si la récompense n'arrive pas au moment où elle est attendue, le système dopaminergique réagit avec une dépression de son activité, au-dessous du niveau de base. Le système reçoit ainsi un signal d'erreur de prédiction sous la forme d'une diminution de l'activité dopaminergique[162]. Il s'agit ici d'un exemple illustrant la capacité du cerveau à encoder un signal d'erreur pour l'utiliser dans une prise de décision, et ceci de manière implicite, sub-personnelle.

Le fonctionnement « prédiction/erreur/mise à jour » n'est cependant pas propre aux neurones dopaminergiques ou à l'apprentissage avec récompense, et son implémentation varie au niveau des différentes régions cérébrales.

Des modèles intégratifs (prédictifs et bayésiens) du fonctionnement cérébral permettent en outre de mettre en évidence l'importance de la confiance dans la prédiction et dans la mise à jour de ses prédictions – donc dans l'apprentissage. Lorsque nous nous sommes confrontés à une nouveauté, nous devons en effet garder un équilibre subtil entre ce que nous avons appris par le passé (par exemple, le constat d'une régularité) et les nouvelles informations, qui peuvent en principe remettre en doute les connaissances préalables. Des études comportementales et en neuro-imagerie indiquent que les réponses données pendant une tâche d'apprentissage (probabilités conditionnelles avec des stimuli perceptifs) répondent à un algorithme optimal d'apprentissage qui prend en compte cet équilibre subtil. Ainsi, une faible confiance dans nos *a priori* nous conduit plus facilement à les mettre à jour en présence de preuves contraires. De même, un *feedback* qui s'associe à cette faible basse confiance influence moins la révision de la prédiction (Meyniel & Dehaene 2017[163] ; Heilbron & Meyniel 2019).

3.3.5 Forces et limites de la confiance métacognitive : une question de calibration

La discussion que nous venons de présenter nous montre que notre architecture cognitive est préparée pour estimer le degré de confiance que suscite une information et que nous pouvons avoir en une opinion ou décision. De plus, cette estimation joue un rôle clé dans des processus fondamentaux comme la décision (au sens large du terme) et l'apprentissage (la mise à jour des *a priori* et des attentes qu'ils produisent). La confiance joue un rôle particulièrement important du fait que nous ne pouvons pas compter systématiquement et tempestivement sur des *feedbacks* correctifs de nos actions : elle permet ainsi de traquer et éventuellement de corriger nos erreurs (Kepecs et al. 2008 ; Yeung & Summerfield 2012 ; Rouault, Dayan & Fleming 2019[164]). La confiance intervient non seulement au niveau personnel, mais également au niveau subpersonnel, y compris dans le contexte de processus de bas niveau. Nous pouvons donc la considérer comme un élément structurel de notre fonctionnement cognitif.
Ce fonctionnement n'est cependant pas sans faille. Les études sur la correspondance entre jugements subjectifs de confiance et performance chez l'adulte (sensibilité métacognitive) mettent en évidence que les jugements métacognitifs de confiance sont souvent meilleurs qu'un choix au hasard, mais tout de même non optimaux, et souvent soumis à erreur et à biais (Metcalfe & Shimamura 1994 ; Dunlosky & Metcalfe 2009 ; Serra & Metcalfe 2009[165]). La dichotomie entre récit métacognitif et performance réelle constitue un objet d'étude en soi, car elle permet de démontrer que décision et confiance répondent à des règles différentes (du moins en partie), ce qui suggère l'existence de substrats neuraux partiellement indépendants (comme mis en évidence dans la discussion sur la confiance et le cerveau) (Grimaldi, Lau & Basso 2015 ; Meyniel, Sigman & Mainen 2015[166]).

Le calibrage de la confiance est défini comme la relation existant entre la moyenne des jugements de confiance (en moyenne, lors des différentes tâches d'un test) et celle des performances. Les jugements de confiance sont considérés comme mal calibrés quand ils divergent de la performance réelle (une confiance élevée dans la performance s'accompagne de réponses incorrectes et une faible confiance s'accompagne de réponses correctes). Nous parlons aussi de précision relative des jugements pour indiquer la capacité d'un jugement à prédire la nature correcte ou incorrecte d'une réponse à un test (Serra & Metcalfe 2009[167]). Certains auteurs invitent en outre à distinguer sensibilité et biais. Concernant ce dernier, lors d'un test, un sujet peut émettre des jugements de confiance globalement calibrés par rapport à ses performances, tout en ayant tendance à juger ses performances de manière plutôt optimiste (Grimaldi, Lau & Basso 2015, voir notamment : Fleming & Lau 2014[168]).

3.3.6 Biais métacognitifs entre sousconfiance et surconfiance

L'un des exemples les plus cités de biais métacognitifs avec calibration négative est celui de la cécité centrale : le sujet résout des tâches visuelles avec un succès supérieur au hasard mais ne s'estime pas capable de le faire. D'autres exemples sont représentés par des formes d'apprentissage implicite, où le sujet commence à donner des réponses correctes avant de se rendre compte qu'il a appris une nouvelle règle (Bechara et al. 1997). Dans la tâche connue comme *Iowa Gambling Task* (Bechara et al. 1997), il s'agit pour des sujets de choisir des paquets de cartes. Le choix se fait initialement au hasard. Ensuite, quand le sujet s'aperçoit de certaines régularités dans les paquets, il choisit sur la base des gains et pertes que les différents paquets apportent. Nous remarquons que les sujets du groupe contrôle (sans lésions préfrontales) commencent à choisir correctement les paquets moins risqués avant de se rendre compte qu'ils le font : pendant une certaine période, ils choisissent donc correctement mais ils pensent choisir au hasard. Leur confiance explicite est mal calibrée, même si un apprentissage statistique a tout de même lieu. (Au contraire, les patients orbito-frontaux n'apprennent pas ou continuent à choisir les paquets désavantageux même une fois qu'ils ont réalisé qu'ils ont plus de chances de perdre.)

Les biais optimistes sont décrits à deux niveaux (Sharot 2011 ; Sharot, Korn & Dolan, 2011 ; Sharot et al. 2012) : au niveau de la cognition de premier ordre, ils se manifestent par la tendance à surestimer la probabilité d'événements positifs et à sous-estimer celle d'événements négatifs ; au second ordre, ils se manifestent par une surconfiance envers les représentations et prédictions ou décisions (croire que nous avons raison, que nos croyances sont vraies, nos décisions correctes[169]). Les effets de surconfiance appartiennent à cette deuxième catégorie – métacognitive –, comme l'effet « Dunning-Kruger » et l'illusion de profondeur explicative.

3.3.6.1 La surconfiance dans les jugements et l'effet dit « Dunning-Kruger »

Plusieurs études mettent en évidence l'existence d'un biais qui affecte les jugements rétrospectifs de confiance (Dunlosky & Metcalfe 2009).

L'effet Dunning-Kruger se révèle lorsque nous mesurons la performance des participants dans des tâches variées, comme l'ironie, la logique, la grammaire… et que nous leur demandons de situer leur performance par rapport aux autres participants : non seulement il existe un effet de surconfiance mais la surconfiance n'est également pas uniforme (Kruger & Dunning 1999[170]). Les études montrent que nous tendons à devenir surconfiants pour des tâches plus difficiles et en situation de sous-performance : plus nous sommes performants, plus nous sommes clairvoyants concernant nos propres performances – voire sous-confiants. Nous sommes donc en présence d'une erreur de prédiction au niveau cognitif (premier ordre) et métacognitif (second ordre), de type optimiste. L'effet Dunning-Kruger a fait l'objet de plusieurs réplications (Dunning et al. 2003 ; Ehrlinger et al. 2008[171]) et il a également été exploré dans le domaine du raisonnement (Pennycook et al. 2017). Les résultats confirment ceux obtenus par Dunning et Kruger.

Dunning et Kruger expliquent que l'effet enregistré est dû au fait que les connaissances qui sont nécessaires pour nous rendre compte que nous manquons de connaissances sont les mêmes que celles qui sont nécessaires pour mener à bien la tâche. Donc, lorsque nous manquons de compétences pour mener à bien une tâche, nous n'avons pas non plus celles nécessaires pour nous en rendre compte : en effet, ces compétences sont les mêmes[172]. L'explication est cohérente avec d'autres observations menées dans le domaine de l'expertise (par exemple, par Chi, Glaser & Rees 1982 sur les compétences en physique). Les études sur l'expertise, même si elles ne portent pas sur la question de l'auto-évaluation des compétences, montrent que les novices, comparés aux experts, manquent de certaines capacités métacognitives en relation avec la tâche, comme la capacité de détecter la difficulté réelle de la tâche, le temps et l'effort nécessaires pour apprendre de leur succès pendant la tâche[173].

Ce type d'étude et les hypothèses explicatives avancées mettent en évidence l'importance de posséder des connaissances « de domaine » pour correctement évaluer la justesse de nos opinions, décisions, performances. Nous reviendrons sur ce point dans la discussion sur les stratégies pour éduquer l'EC, et sur le transfert de compétences. Nous pouvons anticiper en indiquant que les connaissances influent sur les capacités métacognitives – dont celles d'auto-évaluation, la confiance que nous avons dans nos décisions –, ce qui représente une difficulté pour toute approche éducative à l'EC, car l'EC ne peut pas être considéré comme déconnecté des contenus sur lesquels il s'applique.

D'autres caractéristiques de la tâche et de sa formulation influencent cependant les performances. Par exemple, et de manière contre-intuitive, la sur-confiance pourrait être inhibée dans le cas de tâches particulièrement faciles (Lichtenstein, Fischoff & Phillips 1982). Ainsi, alors que des tests difficiles réveillent la sur-confiance, les jugements rétrospectifs dans le cas de tests faciles activent une attitude de sous-confiance. Cet effet s'appelle « l'effet ardu-facile ».

3.3.6.1.1 L'effet Dunning-Kruger dans des tâches de raisonnement

L'effet Dunning-Kruger a également été exploré dans le domaine du raisonnement (Pennycock et al. 2017). Pennycock et ses collègues ont utilisé des tests de raisonnement analytique comme le *Cognitive Reflection Test* et ils ont demandé aux participants d'évaluer leurs performances aux tests et d'indiquer sur une échelle leurs dispositions au raisonnement analytique. Les résultats confirment ceux obtenus par Dunning et Kruger. Pennycock et ses collègues en ont conclu que la capacité de raisonner de façon analytique facilite aussi la capacité métacognitive de se monitorer pendant une tâche de raisonnement[174].

Il semblerait néanmoins qu'un certain degré de sensibilité à l'erreur – même dans le type de raisonnement où nous mettons en évidence des erreurs récurrentes – soit présent au niveau subpersonnel et de manière implicite (Bago & De Neys 2019). De Neys, Rossi & Houdé (2013), De Neys, Moyens & Vansteenwegen (2010), De Neys, Vartanian & Goel (2008) étudient la confiance associée aux performances négatives à des tests de raisonnement - notamment la confiance associée aux performances dans le cadre du problème classique de la batte et de la balle, intégré au CRT développé par Frederick (2005). Les participants donnant la mauvaise (et attendue) réponse au test de la batte et de la balle montrent aussi dans ce test une confiance moins forte en leur réponse que dans d'autres tests témoin. De Neys et ses collègues en déduisent que la détection d'un conflit entre bonne réponse et mauvaise réponse donnée est toujours présente à l'esprit, même si elle est implicite. De Neys, Lubin & Houdé (2014) montrent également que des enfants qui donnent la « mauvaise réponse » dans le cas d'une tâche appelée « de conservation du nombre » ont cependant une confiance moins forte dans leur réponse[175].

Pennycock suggère une possible solution pour le désaccord entre les résultats de son expérience et ceux des expériences de De Neys : même si le signal du conflit est présent au niveau neurologique et peut même être détecté au niveau psychologique, il n'est pas assez « fort » ou efficace pour fléchir la surconfiance. En effet, selon Pennycock et ses collègues, les tests de De Neys montrent que la confiance diminue en relation dans des tests comme celui de la batte et de la balle par rapport à d'autres, mais que la confiance reste néanmoins, incorrectement, élevée (Pennycock et al. 2017[176]).

De Neys, Vartanian & Goel (2008[177]) défendent cependant l'idée que nous sommes – même en relation avec les questions de confiance et de sensibilité à nos biais – plus « intelligents » que nous ne le pensons typiquement, ou que cela n'apparaît pas dans les théories traditionnelles de la rationalité limitée (nous reviendrons sur ce point dans la partie suivante).

3.3.6.2 L'illusion de profondeur explicative

Notre fonctionnement cognitif est dominé, depuis la plus tendre enfance, par des intuitions épistémiques – des intuitions concernant la nature et les lois des phénomènes physiques qui nous entourent, aussi bien que des lois psychologiques, sociales (Rozenblit & Keil 2002[178]). Les bébés découpent le monde (l'information) en objets en suivant certaines « règles » intuitives et spontanées. Ils ne traitent pas tous les objets sur le même plan mais les catégorisent sur la base de certaines propriétés : objets physiques, biologiques, mentaux, sociaux, mathématiques, chaque catégorie recevant son traitement propre et faisant l'objet d'un ensemble d'attentes : les objets physiques ne se mettent pas en mouvement tous seuls, les agents mettent en marche les objets physiques, les objets ne disparaissent pas dans le vide… Ces attentes font appel à des explications. Les enfants sont très actifs dans la recherche (exploration rationnelle) et dans la détection de causes (identification de relations causales à partir d'indices temporels et spatiaux simples). Les signes de ces capacités et attitudes sont présents avant six mois et ont fait l'objet de nombreuses études en psychologie du développement au cours des trente dernières années. La notion de « *core knowledge* » est utilisée pour se référer aux capacités et connaissances plus précoces et ancrées (mais il vaudrait mieux probablement parler de prédispositions à apprendre), de catégories naturelles pour décrire la manière dont se forment les concepts ; et de « théories naïves » pour décrire le fait que adultes et enfants utilisent des intuitions pour s'expliquer les phénomènes qu'ils observent ou pour répondre à des questions qui leur sont posées sur ces phénomènes (parfois en les inventant, mais pas de manière totalement aléatoire) (Spelke & Kinzler 2007).

Cependant, nous ignorons les lacunes des explications que nous fournissons. Nos connaissances sont limitées mais nous nous faisons l'illusion d'en savoir plus qu'en réalité, et notamment de comprendre plus qu'en réalité les mécanismes qui sous-tendent le fonctionnement d'objets d'usage commun (Sloman & Fernbach 2018). Rozenblit & Keil (2002) étudient ce phénomène chez l'adulte. Leur méthodologie repose sur des jugements de confiance : les participants évaluent sur une échelle de 1 à 7 leur compréhension d'un phénomène ; ensuite, ils décrivent le fonctionnement causal de l'objet observé dans le détail et réévaluent leur compréhension sur la même échelle ; ils répondent à une question diagnostique précise sur l'objet et ils se réévaluent ; ils lisent un texte explicatif concernant l'objet et ils se réévaluent, etc. Enfin, ils répondent à une question concernant l'utilité du texte explicatif fourni pour comprendre le phénomène. Les résultats indiquent que les sujet surestiment initialement leur compréhension et que leur confiance diminue suite à l'effort pour fournir une explication détaillée. Dans la phase de débriefing, les participants à l'étude déclarent leur surprise relative à leurs pauvres connaissances. L'étude a été ensuite réalisée par les mêmes auteurs avec d'autres populations (mais encore homogènes) et en variant des paramètres. Par exemple, dans un test, les participants ont été informés dès le début qu'ils auraient à fournir une explication détaillée du mécanisme de fonctionnement de l'objet observé : cela a réduit la différence entre auto-évaluations successives, mais ne l'a pas fait disparaître. Les auteurs ont aussi testé d'autres formes de connaissances et de compréhension avec le même type de dispositif : connaissances factuelles, procédurales, narratives (le résumé d'un film). Ils ont remarqué des différences significatives : surconfiance pour les faits, mais inférieure par rapport aux mécanismes, et bon calibrage pour les procédurales et narratives. En revanche, les mêmes résultats sont obtenus pour les mécanismes et pour des explications de phénomènes naturels. Les explications causales seraient donc particulièrement susceptibles de susciter de la surconfiance par rapport à d'autres contenus de connaissance. Ce qui signifierait que la surconfiance n'est pas un phénomène général mais qu'elle est relative à certains domaines de connaissance[179].

Rozenblit & Keil (2002) ne considèrent pas cette illusion comme identique à l'illusion optimiste qui nous amène à être surconfiants en nos connaissances et capacités – un phénomène largement documenté dans des domaines assez divers. L'illusion de profondeur explicative traduit en effet une surconfiance spécifique dans la compréhension de mécanismes de type causal, donc dans le cadre d'explications. Nous avons l'illusion de posséder plus de

connaissances que nous n'en avons réellement, et de mieux comprendre les mécanismes qui sous-tendent le fonctionnement d'objets d'usage commun que nous le faisons en réalité. Cependant, les mêmes personnes qui sont surconfiantes quand il s'agit de la justesse de leurs explications causales ne le sont pas face à d'autres formes de connaissance, notamment des connaissances factuelles ou procédurales. Plusieurs raisons peuvent entretenir ce type d'illusion. Nous pouvons avoir l'impression de comprendre les relations complexes entre composantes d'un mécanisme parce que nous en comprenons le fonctionnement global. Le fait de posséder des explications même lacunaires peut représenter une utilité immédiate et entretenir l'impression d'utilité sur le long terme, d'où le terme d'illusion de profondeur explicative. D'autres raisons se situent également au niveau du *feedback*. Par exemple, les explications causales concernant les mécanismes de fonctionnement d'un objet que nous manipulons obtiennent un *feedback* positif à partir du moment où nous savons mettre en marche l'objet en question, mais ce *feedback* est trompeur par rapport à la nature lacuneuse de notre compréhension profonde de ses mécanismes de fonctionnement. De plus, des explications en profondeur sont rarement demandées et donc nous manquons de *feedback* à ce niveau (Rozenblit & Keil 2002[180]).

3.3.7 Pourquoi la sensibilité métacognitive pourrait-elle être mal calibrée ?

Plusieurs réponses existent concernant la question du mauvais calibrage de la confiance, lorsqu'elle a lieu.

Premièrement, les biais optimistes et la surconfiance pourraient avoir une valeur adaptative[181]. Celle-ci contrebalancerait la valeur adaptative négative de l'ignorance de risques réels, de l'optimisme excessif et de la confiance excessive en soi qui peuvent amener à des conséquences désastreuses. Sharot & Garrett (2016) remarquent par exemple que le rôle de la valeur positive du *feedback* concerne notamment les cas des croyances motivées et de l'information ambigüe. Selon eux, la tendance à donner une interprétation optimiste aux informations est présente chez différentes espèces (oiseaux, cochons) et elle se manifeste notamment lorsque la situation est ambigüe et incertaine et en absence de stress (Sharot & Garrett 2016).

Deuxièmement, la prédiction de réussite dépend aussi de l'expérience préalable avec des tâches semblables. En l'absence de cette expérience, l'enfant, et plus généralement le débutant, peut se montrer surconfiant (Proust 2019[182] ; Lockl & Schneider 2007).

Troisièmement, le fait que les sentiments métacognitifs soient le produit d'heuristiques basées sur des indices indirects de la validité de la réponse peut créer des pièges qui en affectent la fiabilité (Serra & Metcalfe 2009[183] ; Koriat 2012). Quels sont donc les indices desquels naissent les sentiments métacognitifs ?

3.3.7.1 Les heuristiques à l'origine des illusions métacognitives : le problème de la nature indirecte des indices

Serra & Metcalfe (2009) identifient plusieurs heuristiques à l'origine des illusions métacognitives : la familiarité avec un certain matériel peut influencer la sensation de bien le connaître[184] ; la rapidité avec laquelle un souvenir est récupéré dans la mémoire peut influencer la confiance dans les réponses et d'autres jugements liés à la capacité de se souvenir et d'apprendre[185].

Koriat (2012) en particulier propose un modèle du jugement de confiance basé sur la disponibilité en mémoire. Dans ce modèle, la confiance dans une réponse est corrélée avec des indices comme la rapidité avec laquelle la réponse est retrouvée dans la mémoire ou sélectionnée ; elle est aussi corrélée avec la justesse de la réponse, mais uniquement de façon

indirecte, *via* l'information sur la rapidité. En d'autres termes, la rapidité de récupération (*ease of retrieval*) est un indice frugal et indirect pour juger de la justesse de sa réponse[186]. Les associations qui sont les premières activées par la tâche et la réponse dépendent des expériences préalables et de lois de l'apprentissage qui font que certaines associations sont apprises ou réactivées plus facilement que d'autres. La rapidité peut donc être un bon indice, notamment lorsque le sujet possède beaucoup de connaissances dans le domaine de la tâche ; mais, en tant qu'indice indirect, elle peut aussi conduire le jugement métacognitif en erreur.

Un autre indice utilisé pour juger, implicitement et de façon frugale, de la justesse de la réponse est la cohérence (*self-consistency*) entre différentes représentations présentes dans la mémoire. Nous pouvons voir la chose de la manière suivante : le système de confiance ne cherche pas à mesurer directement la validité de la réponse (*validity*, un jugement qui lui serait inaccessible) mais la probabilité de donner la même réponse dans le futur (*reliability*). Cette dernière constitue à son tour un indice de la validité de la réponse (*validity*). Or, la tâche active une certaine réponse et une série de représentations. Chaque représentation a un signe positif ou négatif en relation avec la réponse donnée. Si la somme est positive, le système de confiance détecte un état de cohérence et produit une sensation de confiance dans la réponse choisie. Un mauvais calibrage de la confiance peut se produire en vertu de la nature indirecte de l'indice[187]. Plusieurs aspects de la situation d'apprentissage peuvent en outre produire différentes formes d'illusion de connaissance[188].

3.3.8 Stratégies qui permettent d'améliorer la qualité des jugements métacognitifs

En dépit de l'existence de mécanismes au niveau neural et psychologique pour détecter les erreurs et conflits – à la lumière des biais que nous avons décrit, et notamment de l'effet Dunning Kruger dans ses différentes manifestations, nous sommes confrontés à un problème assez profond pour le développement de l'EC – conçu comme la capacité d'évaluer correctement la qualité de l'information en notre possession pour décider d'un cours d'action. L'évaluation des informations est en effet d'un côté bien servie par des mécanismes naturels et « profonds » (en relation avec le niveau de traitement de l'information) mais ces mécanismes peuvent avoir du mal à se traduire dans des évaluations calibrées, lorsque nous passons au niveau conscient, en « première personne ». Existe-t-il des solutions ? Ont-elles été évaluées expérimentalement ?

Même s'ils sont naturels, précoces et silencieux, les sentiments métacognitifs comme la sensibilité à l'erreur peuvent être cultivés.

D'une part, une stratégie qui semble pouvoir influer positivement sur la précision des jugements métacognitifs et sur leur calibrage consiste en la prise de conscience des biais qui influencent ces mêmes jugements. Il s'agit notamment de prendre conscience des heuristiques et indices à la base des jugements de confiance, et de comprendre comment ceux-ci peuvent produire des illusions de compréhension, d'apprentissage ou d'avoir donné la réponse correcte. L'effet des tests avec *feedback* améliorerait en effet les performances mais sans nécessairement produire un transfert sur du nouveau matériel ou sur de nouveaux contextes. L'explicitation des raisons pour lesquelles nous pouvons faillir aurait cet effet supplémentaire (Serra & Metcalfe 2009[189]).

Butterfield & Metcalfe (2006) utilisent des questions de type « Trivial Pursuit » pour évaluer la capacité de revenir sur une réponse qui s'accompagne d'un degré élevé de confiance. Après avoir répondu aux questions, les participants estiment leur confiance dans chaque réponse qu'ils ont donnée. Leurs réponses sont corrigées par un *feedback* qui leur signale lesquelles étaient correctes. Puis, les mêmes sujets sont re-testés sur les mêmes questions quelque temps plus tard. Butterfield & Metcalfe (2006) montrent, en premier lieu, que la présence d'un *feedback* permet de changer de réponse, même dans les cas où la confiance dans la réponse avait été très élevée. Ensuite, que les questions sur lesquelles les participants se sont trompés,

et pour lesquelles ils se sentaient particulièrement confiants, sont celles qui, lors du deuxième test, sont corrigées avec la plus grande probabilité. il s'agit de l'effet « d'hypercorrection ». Cet effet est contre-intuitif, car nous nous attendrions à voir plus difficilement changer les positions sur lesquelles les participants se sentaient plus confiants.

Selon les auteurs, plus grande est l'erreur de confiance, plus grand est l'effet surprise, et donc plus grande est aussi l'attention portée à l'erreur. Cette explication est testée en proposant un stimulus auditif en même temps qu'est donné le *feedback*. La perception du stimulus est moins probable dans le cas du couple *feedback*-haute confiance. L'attention est donc dans ce cas plutôt capturée par le *feedback*, d'où la plus grande mémorisation du *feedback* et la plus grande probabilité de corriger l'erreur en situation de surconfiance (hypercorrection)[190].

Avec un dispositif semblable, Butler, Karpicke & Roediger (2008) testent l'effet du *feedback* sur les réponses correctes mais accompagnées d'un niveau de confiance faible. Ils proposent des tests avec réponses à choix multiples. Les sujets sont forcés de répondre. Cela donne lieu à des situations où les sujets peuvent donner des réponses correctes par hasard, mais sans se sentir réellement compétents. Ces tests montrent que, à partir d'une réponse correcte mais accompagnée d'un faible niveau de confiance, le *feedback* améliore la performance cognitive de premier ordre (le fait de savoir donner des réponses correctes au deuxième test) probablement en passant par une amélioration de la capacité métacognitive[191].

Dans une expérience impliquant environ 60 sujets, Carpenter et al. (2019) comparent l'effet du *feedback* sur la justesse de la réponse et celui du *feedback* sur la justesse des jugements métacognitifs sur la performance. La motivation pour essayer d'influencer directement la confiance métacognitive sans chercher à influer sur les performances cognitives de premier ordre est double : d'un côté, les auteurs reportent des résultats positifs récents d'amélioration des performances métacognitives obtenus *via* des interventions telles que la médiation, l'absorption de substances ou la stimulation cérébrale ; de l'autre, ils expliquent que la recherche en neurosciences a mis en évidence l'existence de substrats neuraux spécifiques pour les performances métacognitives, séparés des substrats impliqués dans les performances cognitives de premier ordre (Carpenter et al. 2019[192]). Nous avons cité plus haut les résultats positifs des interventions qui se proposent d'influencer directement le calibrage de la confiance *via* des méthodes explicites. À différence de celles-ci, l'intervention testée *via* la plateforme *Amazon Mechanical Turk* par Carpenter et al. (2019) est de type implicite, car elle se base uniquement sur la proposition de *feedback*. L'expérience comprend une tâche de décision perceptive et une tâche de mémoire. Dans la tâche perceptive, les participants jugent la luminosité relative de deux formes abstraites qui leurs sont proposées simultanément. La tâche de mémoire implique elle la visualisation rapide d'une série de lettres et la proposition de deux choix alternatifs pour reconnaître l'image initiale. Suite à leur choix, les participants effectuent un jugement rétrospectif de leur performance, en déclarant leur confiance dans la réponse sur une échelle allant de 1 (confiance très basse) à 5 (très haute confiance). Après 27 essais, les participants reçoivent un *feedback* concernant respectivement les performances (groupe témoin) et leur prestation métacognitive (groupe expérimental). Ils continuent de la même manière pendant dix sessions d'entraînement. Les résultats indiquent que, juste après la première session, les participants du groupe expérimental voient diminuer leur biais métacognitif, alors que l'efficacité de la sensibilité métacognitive augmente progressivement mais plus lentement au cours des entraînements. Le fait de recevoir un *feedback* sur la performance de premier ordre ne produit pas le même effet. La proposition alternée de tâches perceptives et de mémoire permet d'évaluer si l'amélioration des capacités métacognitives dans l'un des deux domaines se généralise dans l'autre. Les résultats indiquent qu'un transfert a lieu de l'entraînement perceptif aux performances en métamémoire, mais pas pour la composante de la sensibilité métacognitive. L'expérience conduite par Carpenter et al. (2019) ne permet pas de savoir si ce type de gain se transfère également à des tâches de la vie réelle, ni quelle serait la durée du bénéfice obtenu, mais ouvre la voie à d'autres études complémentaires.

Selon la nature des stimuli (assez simples) utilisés dans l'étude, nous pourrions faire l'hypothèse que les effets de l'entraînement sont d'autant plus forts et susceptibles de transfert s'ils sont accompagnés d'instructions concernant les heuristiques qui influencent les performances métacognitives et si l'entraînement est mené sur des tâches plus réalistes.

La question de la nature de la métacognition – dépendante ou indépendante du contexte – ne semble pas être encore résolue (Rouault et al. 2018[193]).

3.3.9 La confiance métacognitive. Conséquences pour l'éducation de l'EC

L'EC consiste – selon la définition que nous en avons donnée – à évaluer le plus correctement possible la qualité d'une information afin de prendre une décision. Cependant, cette évaluation ne comporte pas uniquement une accumulation d'indices sur le fait qu'une information en notre possession est fiable ou ne l'est pas. L'évaluation donne également lieu à un état, sentiment ou jugement explicite de confiance en cette information. Cette dernière influe, à son tour, sur nos prises de décision et opinions présentes et futures.

La confiance métacognitive (sensibilité à l'erreur et confiance implicite) apparaît en effet comme un mécanisme fondamental pour prendre en compte la valeur d'une information afin de prendre une décision, réguler la capacité à changer d'opinion sur la base de nouvelles informations et mettre à jour ses représentations et prédictions. Son fonctionnement est actif précocement dans la vie de l'enfant, mais, à l'état naturel, les jugements de confiance – implicites et explicites – s'appuient sur des indices indirects et sur des heuristiques. Cette dépendance peut donner lieu à des erreurs de jugement.
De ces considérations émergent des indications pour l'éducation de l'esprit critique. L'indication principale consiste à outiller la confiance à porter à une certaine information ou opinion. La confiance métacognitive peut en effet faire l'objet d'apprentissages, et différentes méthodes sont potentiellement utiles. La littérature concernant métacognition et confiance fournit ainsi non seulement un cadre théorique mais aussi des indications pratiques concernant la manière d'améliorer ces capacités qui ne sont pas sans limites et qui sont affectées par des illusions ou biais comme les autres capacités.

La littérature analysée nous a permis d'identifier les facteurs qui influencent positivement le calibrage efficace de la confiance :

- le fait de posséder des connaissances dans le domaine concerné. La connaissance de domaine semble être une condition pour le bon exercice de la confiance dans les informations en notre possession, dans nos prévisions et décisions ;
- la présence de *feedback* appropriés. Le *feedback* peut concerner la performance, mais il peut aussi être utile de fournir un *feedback* concernant le bon ou mauvais calibrage de la confiance et la présence d'un biais positif ou négatif de confiance, c'est-à-dire d'informer le sujet sur la distance entre ses jugements de confiance et la réalité. Pour arriver à cela, la confiance doit être évaluée puis comparée aux performances. Elle doit donc faire partie de la pédagogie de l'esprit critique. Le fait d'inclure la métacognition dans les apprentissages de manière assumée, et de se doter d'outils comme les tests avec *feedback*, a d'ailleurs montré son utilité dans d'autres domaines de l'apprentissage. Son introduction du côté de l'éducation de l'esprit critique ne saurait donc qu'être cohérent avec des indications données par ailleurs aux éducateurs (voir en particulier : Proust 2009 ; Hattie, Biggs & Purdie 1996 ; Higgins et al. 2012) ;
- la confiance est associée à la performance, mais elle l'est de façon indirecte, *via* des indices. Une indication consiste à faire de ces critères et des illusions potentielles de jugement métacognitif l'objet d'une instruction explicite, afin de les identifier et de les

anticiper. L'explicitation semble représenter un élément clé du transfert, même si d'autres stratégies comme le *feedback* dument administré sont mises en place ;

- comme le recours à des indices indirects pour juger la confiance que devrions donner à une opinion ou décision est soumis à erreur, une stratégie consiste de se doter de critères plus objectifs, externes. La littérature sur la métacognition met en évidence que des aides externes telles que les évaluations et les tests peuvent être nécessaires afin de corriger les illusions métacognitives de connaissance et d'apprentissage. Cette indication peut être exportée à la sensibilité métacognitive de la manière suivante : parfois, nos métriques internes sont insuffisantes pour nous fournir une estimation correcte de la confiance que mérite une information ou une représentation que nous entretenons. Dans ce cas, nous voulons adopter des échafaudages externes analogues à ceux rencontrés dans la première partie : des grilles de critères qui permettent d'objectiver la qualité d'une information ou d'une autre représentation. La confiance subjective doit alors se plier à ces critères, que le sujet a reconnu comme supérieurs à ses indices naturels en vertu d'une instruction explicite sur ces derniers.

Ces outils sont d'autant plus nécessaires quand nous devons évaluer nos propres opinions, positions et décisions. Nous avons en effet constaté qu'il est difficile d'évaluer sa propre compétence et expertise. Pourtant celle-ci est une condition nécessaire pour savoir si nous pouvons nous faire confiance au moment de prononcer un jugement sur n'importe quelle information. Nous pouvons alors supposer que la connaissance des biais cognitifs et métacognitifs représente un outil de plus, nous permettant de mieux anticiper nos chances d'erreur, d'identifier les situations plus à même de nous mettre en difficulté en créant des obstacles à un jugement juste de confiance ou de fiabilité dans une information.

Pour conclure cette discussion sur la métacognition, il est important de souligner que l'objectif d'outiller les mécanismes naturels de confiance consiste à atteindre une confiance mieux calibrée, une meilleure sensibilité. Cela se traduit par la capacité à se déclarer incertain quand les preuves à notre disposition sont de faible qualité, ou manquent. Par exemple, quand les seules preuves présentes ne sont que des anecdotes personnelles, nous devrions légitimement baisser notre confiance dans l'opinion ou dans l'information concernées. Ceci ne signifie pas que nous nous retiendrons de former une opinion, car cela serait irréaliste, mais que nous devrions juger l'opinion comme faiblement soutenue, et donc comme susceptible d'être renversée, modifiée, à la lumière de nouvelles preuves. Lorsque les preuves sont solides, et que nous jugeons correctement que nous savons les évaluer à leur juste valeur, grâce à nos connaissances, alors notre confiance est légitimement élevée, et nous pouvons nous déclarer certains de notre position.

Nous pouvons représenter cette manière de concevoir la confiance bien calibrée sous la forme d'un continuum allant de « pas confiant du tout » à « très certain ». L'exercice de l'EC consiste à se placer au bon endroit sur la ligne, sur la base des informations à notre disposition, y compris l'évaluation de notre propre compétence et expertise pour évaluer ces informations.

Pourquoi expliciter cela est-il important ? Parce que, souvent, l'EC est plutôt décrit en termes de méfiance et associé à une image défensive, protective, de doute ou de vigilance (comme le terme « vigilance épistémique »). Or, la discussion autour de la confiance met en évidence l'importance d'être confiants et méfiants à bon escient, donc de calibrer cette confiance sur la base d'une évaluation des preuves à l'appui, des sources, de la plausibilité et de la pertinence de l'information, de notre propre compétence. Au fur et à mesure, cette évaluation devient plus sophistiquée, plus adaptée au contexte, grâce aux outils artificiels qui s'adressent aux capacités naturelles ; la confiance devient elle aussi mieux calibrée.

3.4 Chercher à comprendre la nature de nos erreurs

C'est un fait que nous commettons des erreurs dans le processus qui permet de prendre des décisions ou de se forger des opinions. Nous avons vu que ces erreurs peuvent affecter de manière spécifique l'évaluation des informations (erreurs de jugement des contenus informationnels ou des sources de l'information) et le calibrage de notre confiance (confiance excessive en une décision que nous aurions prise, en une prise de position, opinion, perception ; ou confiance trop basse, alors que nos informations sont de bonne qualité). De plus, nos décisions, nos jugements, nos évaluations, nos perceptions sont le produit d'inférences silencieuses qui ont lieu au-dessous du seuil de la conscience, et qui échappent à la réflexion.

Souvent, l'attention est focalisée sur la manière dont ces processus implicites nous mènent en erreur. Ainsi, une vaste littérature en psychologie de la décision et du raisonnement est dédiée à la description des biais qui influencent nos choix, nos jugements et nos opinions (voir par exemple, Pohl 2017 ; Kahneman 2011 ; Gilovich, Griffin & Kahneman 2002 ; Ariely 2008). Cette littérature étend son succès au-delà de son domaine propre et est souvent évoquée dans le domaine de l'EC pour expliquer nos erreurs les plus communes de perception, de prévision, d'explication, et plus généralement de décision (voir, à titre d'exemple, Battersby & Bailin 2013 ; Halpern 2013 ; Nisbett 2015 ; Stanovich & Stanovich 2010).

Cette littérature appartient notamment au programme dit des « heuristiques et biais » en psychologie sociale et du raisonnement[194]. L'idée qui sous-tend ce programme est la suivante : dans des conditions de jugement en situation d'incertitude, nous avons recours à des solutions-raccourcis qui sont en nombre limité, rapides à se mettre en place et simples (Tversky & Kahneman 1974). Ces heuristiques sont généralement utiles mais peuvent parfois nous induire en erreur. Quand une heuristique donne lieu à une erreur systématique, nous parlons de biais (Kahneman & Tversky 1996). La grande variété des biais observés dans le raisonnement pourrait être reconduite à un petit nombre d'heuristiques fonctionnelles (Kahneman et al. 1982 ; Gilovich, Griffin & Kahneman 2002).

3.4.1 Le programme « heuristiques et biais » (H & B) et la nature des erreurs

Le programme « H & B » puise ses racines dans la notion de rationalité limitée (*bounded rationality*) introduite par Herbert Simon pour décrire les effets des limites computationnelles de notre architecture cognitive sur nos choix et décisions. Sa base théorique est donc constituée par l'idée selon laquelle notre cognition est intrinsèquement limitée et que des solutions efficientes peuvent être développées pour réduire les coûts computationnels et mener à des décisions satisfaisantes, non nécessairement optimales[195]. Les biais cognitifs seraient une manifestation de telles solutions. La valeur ajoutée d'étudier les biais cognitifs pour le psychologue consiste à gagner accès aux solutions implémentées de manière structurelle dans notre architecture cognitive. Pour le système cognitif qui les implémente, ces solutions permettent d'atteindre un certain niveau d'efficience, puisqu'elles ne prennent pas en compte les solutions optimales que nous pouvons trouver en prenant le temps et en utilisant des algorithmes de raisonnement plus sophistiqués et adaptés à la situation ; elles sont aussi sources d'erreur.

Ce programme a été reçu de manière très positive par une grande partie des chercheurs en psychologie du raisonnement et jouit d'un large succès dans des disciplines diverses (médecine, économie, loi, management, politique) et auprès du grand public. (Il suffit de penser au succès de livres comme *Thinking Fast and Slow* de Daniel Kahneman, l'un des porte-paroles de ce programme plus connu par le grand public : Kahnemann 2011).

3.4.2 Critiques du programme « H & B »

Le programme a également été critiqué, de manière plus ou moins sévère. Les critiques principales à l'analyse des heuristiques et biais viennent d'approches à composante plus écologique, évolutionnaire, qui ont donné lieu à des programmes alternatifs, comme celui de la rationalité adaptative (Cosmides & Tooby 1994 ; Gigerenzer 1991, 1996, 2004 ; Gigerenzer, Hoffrage & Goldstein 2008 ; Haselton et al. 2009). Ce programme ne se limite pas à la cognition humaine, mais nous le retrouvons aussi dans le domaine de l'écologie et de la cognition animale, où la rationalité adaptative des décisions est souvent étudiée relativement à des tâches de recherche de nourriture, voir Stephens 2008).

Parmi les critiques les plus importants du programme H & B se trouve le psychologue Gerd Gigerenzer. Il soutient notamment que la manière dont les questions sont posées par les psychologues peut créer des pièges dans lesquels notre cognition tombe en produisant une erreur, alors que nous ne nous tromperions pas si la question était posée autrement (Gigerenzer, Hoffrage & Goldstein 2008). En outre, trop d'attention aurait été portée aux cas dans lesquels le système cognitif est mis en échec et amené à commettre des erreurs, ou juste rapide mais peu satisfaisant d'un point de vue rationnel, et trop peu d'attention aurait été portée aux situations dans lesquelles notre cognition nous sert bien, voire de manière optimale[196]. En contraste avec les biais en tant qu'effets de solutions « *quick and dirty* », Gigerenzer, Tood & The ABC Research group (1999) proposent donc un ensemble de heuristiques qui sont plutôt « *fast and frugal* » : des heuristiques qui nous rendent plus efficaces et efficientes, qui ont un rôle adaptatif, et qui font partie d'une boîte à outils mentale ayant été sélectionnée au cours de notre évolution en réponse à des tâches et problèmes spécifiques à résoudre. La prise de décision implique de choisir l'outil le plus approprié dans la boîte à outils. Un exemple est constitué par l'heuristique de reconnaissance : imaginons que l'on nous demande laquelle parmi deux villes a le plus d'habitants, que nous ne connaissions pas la réponse mais que nous ayons à nous baser sur des critères. Celui que nous utiliserions probablement serait celui de reconnaissance (heuristique de reconnaissance) : nous reconnaissons le nom de l'une des deux villes donc nous répondons que c'est la plus grande. Le critère en question nous sert bien dans ce cas, car il est en effet relié à la taille de la ville : des villes plus grandes sont plus souvent nommées à la télévision, etc. Cependant, comme tous les critères indirects, celui-ci peut ne pas s'appliquer dans certains cas, ou être « activé » pas de mauvaises raisons.

Les approches évolutionnistes adaptatives reprochent à la vision « heuristiques et biais » d'avoir recours uniquement à des **normes externes** pour juger si un comportement est rationnel ou irrationnel. Dans l'approche H & B, la norme est établie sur la base de la désirabilité d'un certain type de comportement selon un point de vue externe, logique. Alors que, dans l'approche adaptative, nous cherchons à établir si la norme correspond réellement aux objectifs auxquels le système cognitif répond afin d'augmenter la valeur sélective de l'individu, dans des conditions données. L'approche évolutionnaire prend donc en compte les conditions écologiques dans lesquelles un certain comportement émerge[197]. Par exemple, si la recherche de la vérité est une norme importante du point de vue culturel, rien n'assure que les solutions sélectionnées au cours de l'évolution répondent uniquement à de telles normes. Cette sélection s'est faite sous des contraintes multiples – dont la recherche de vérité, mais pas seulement – et dans des environnements complexes, que nous négligeons souvent dans une approche H & B classique.

3.4.3 L'approche « évo-éco » : remettre nos intuitions et instincts dans une lumière positive

Un concept fondamental de l'approche évolutionniste et écologique (« évo-éco ») est celui de *match* (et de *mismatch*) entre les solutions prêtes à l'emploi (heuristiques) que nous mettons

en œuvre et l'environnement dans lequel nous agissons dans le présent et le passé. Les solutions disponibles dans notre boîte à outils cognitive sont le fruit de notre évolution ; leur structure et fonctionnement sont donc déterminés par les conditions de notre passé où leur évolution a eu lieu (contraintes écologiques). Une certaine heuristique n'est ni rationnelle ni irrationnelle en soi : elle est adaptée à un certain contexte et donne une réponse optimale dans ce contexte. La sous-optimalité apparente d'une stratégie peut alors s'expliquer de différentes manières : il est possible qu'il s'agisse d'une limite physiologique et que l'outil pourrait ou aurait pu devenir plus performant, mais d'autres causes peuvent être envisagées. Par exemple, l'outil était optimal dans le contexte passé dans lequel il a été sélectionné ; il est le meilleur compromis quand nous considérons l'ensemble des contraintes qui agissent sur lui, et non simplement celle qui nous intéresse immédiatement. Si nous voulons bien juger une heuristique (ou tout trait biologique), nous devons donc imaginer le contexte original de son utilisation[198].

3.4.3.1 Un exemple d'une approche « évo-éco » à l'explication de décisions et comportements

Pour mieux comprendre le changement de perspective qu'implique une approche évolutionniste sur l'évaluation d'un choix de stratégie (prise de décision), nous allons nous éloigner quelques instants des « préoccupations humaines » liées à la prise de décision et à la rationalité pour nous tourner vers des études comportementales concernant d'autres espèces. Cette approche « comparative » permet d'émettre des hypothèses concernant des comportements humains « irrationnels » en apparence.
En écologie comportementale, des modèles dits « *Optimality Foraging Theory* » (théorie de la recherche optimale de nourriture) permettent de déterminer les contraintes qui agissent sur un animal quand il foit réaliser des choix pour la recherche de nourriture. Cette approche permet de mieux caractériser la valeur adaptative d'un comportement. Ainsi, des études menées sur des corbeaux (Richardson & Verbeek 1986) montrent que ces animaux laissent une partie de leurs proies (des palourdes) intactes après les avoir localisées. Pourquoi cette nourriture potentielle est-elle délaissée ? Est-ce une perte d'énergie consécutive à un comportement irrationnel ? Les corbeaux acceptent 100 % des proies d'une certaine taille et la proportion diminue en dessous. En considérant les bénéfices caloriques des proies de différentes tailles et les coûts pour rechercher, creuser, ouvrir la proie et se nourrir, les chercheurs ont construit un modèle mathématique basé sur l'hypothèse selon laquelle les oiseaux maximiseraient les calories absorbées : un régime optimal. Les observations correspondent aux prédictions du modèle. Les chercheurs ont par ailleurs établi un lien entre la valeur sélective et le gain d'énergie acquise par jour. Cela valide donc l'approche : en supposant que le comportement était adaptatif, et sans *a priori* trop fort sur les contraintes évolutives, ils ont révélé celles qui s'exerçaient réellement sur le choix de l'animal et déterminé que son comportement était réellement optimal. D'autres fois, les prédictions du modèle s'écartent de la réalité observée. Si d'autres facteurs que la prise de nourriture influencent le choix des proies, le modèle de maximisation calorique est alors rejeté après le test. Il est ensuite possible d'ajouter d'autres paramètres, par exemple la présence de prédateurs. Certaines contraintes peuvent aussi s'appliquer sur une partie de la population seulement.

Les modèles écologiques d'optimalité ne cherchent donc pas à détecter des comportements irrationnels par rapport à des normes économiques préétablies. Le postulat considère que la sélection naturelle a optimisé des comportements sur la base de leur valeur sélective. La stratégie consiste donc à établir les contraintes ayant façonné ces comportements aujourd'hui observés (Stephens 2008). Conclure à l'« irrationalité » d'un comportement dans un cadre expérimental donné implique donc de comprendre au préalable les causes distales du comportement en question.

Cette approche peut être importée dans la problématique du raisonnement et de la prise de décision humains. c ainsi que Gigerenzer (1991) supposent par exemple que le fait d'échouer

aux tests d'évaluation de la prévalence d'une maladie (négligence du taux de base – Casscells, Schoenberg & Grayboys 1978) est lié au fait que les participants doivent évaluer la probabilité d'un événement unique : « Certains de nos mécanismes de raisonnement incluent des aspects du calcul des probabilités mais ils sont conçus pour travailler à partir d'une information sur la fréquence pour produire des fréquences. » La reformulation du problème de Casscells en termes de fréquences et imposant une réponse en termes de fréquence augmente significativement les performances (76 % contre 18 %). D'où la nécessité de considérer les modalités de raisonnement sous l'angle adaptatif en se demandant : quels problèmes les humains ont-ils eu à affronter au cours de leur histoire évolutive ? Pourquoi, par exemple, n'ont-ils pas évolué pour traiter ces probabilités de cas individuels ? Il se peut que ce soit parce que les estimations fondées sur les fréquences sont utiles pour prendre des décisions. Gigerenzer (1991) avance qu'il est impossible d'évaluer la probabilité d'un événement isolé : il est très difficile d'évaluer la classe de référence : quelles sont nos chances d'être frappés par la foudre : tout dépend de notre classe de référence, si nous faisons beaucoup de randonnées ou que si nous ne sommes jamais dehors, etc. De plus le nombre de classes de référence est potentiellement infini. Alors que nos ancêtres pouvaient tirer un bénéfice du développement de mécanismes sensibles à la fréquence : « Si nous avons le choix entre deux parcelles, nous utilisons celle qui a bien rendu auparavant. »

Sur la base des considérations exprimées ci-dessus, d'autres tendances considérées comme menant à des prises de décisions irrationnelles – l'aversion au risque par exemple – peuvent être réinterprétées comme des stratégies très efficaces selon que nous considérons une optimalité plutôt qu'une autre (par exemple, minimiser la variance des gains – sauf si des gains forts sont nécessaires – plutôt que simplement maximiser les gains) (Cosmides 1989 ; Cosmides & Tooby 1996 ; Gigerenzer 1991 ; Gigerenzer & Hug 1992 ; Oaksford et Chatter 1994 ; Rode et al. 1999).

Prenons l'exemple suivant : en proposant deux boîtes aux participants, l'une contenant 50 boules noires et 50 boules blanches, l'autre contenant 100 boules noires et blanches de répartition inconnue, ils choisissent consécutivement deux fois la première boîte pour trouver une boule noire (supposant donc que la probabilité de trouver une boule blanche est plus faible) puis ensuite une boule blanche (Rode et al. 1999). Cela est interprétée en termes d'évitement du risque (toujours prendre la boîte minimisant le risque), contre tout raisonnement naturel (il n'est pas possible que la boîte A soit à la fois favorable pour trouver une boule noire et une boule blanche puisque les hypothèses sont mutuellement exclusives). Selon les auteurs, il est plus probable en utilisant cette stratégie de gagner deux fois consécutivement (1/4) mais moins probable de ne pas perdre deux fois. Un besoin égal à une victoire doit nous pousser vers la stratégie « risquée » alors qu'un besoin égal à deux victoires doit nous pousser vers la stratégie « évitement du risque ». Ce type d'approche expérimentale montre que les participants possèdent des stratégies rationnelles, non du point de vue mathématique ou logique, mais du point de vue écologique : bien adaptées pour la résolution de problèmes adaptatifs, auxquels ont fait face nos ancêtres durant notre histoire évolutive (minimiser la variance sauf en cas de besoin élevé lors de la recherche de fruits dans des situations incertaines). Ce que nous considérons comme un biais menant à des comportements irrationnels peut en réalité être complètement réinterprété comme une stratégie « *better than rational* ».

3.4.3.2 Un exemple d'approche « évo-éco » pour l'explication des illusions perceptives

La notion d'erreur a fait l'objet de critiques même dans des domaines différents que celui du raisonnement. Les illusions cognitives, ou biais, sont en effet souvent définies par analogie avec les illusions perceptives comme étant des erreurs (déviations de la réalité) systématiques (à la fois prédictibles et robustes au niveau inter et intrapersonnel), involontaires et difficiles voire impossibles à contrôler, surprenantes. Les illusions cognitives sont considérées comme plus problématiques à définir par rapport aux perceptives notamment au niveau de la notion

d'erreur. En effet, il serait plus difficile d'établir la présence d'une erreur car la « norme » contre laquelle l'erreur est définie est moins objective. Les illusions cognitives seraient aussi moins « inflexibles » que celles perceptives, du moins certaines d'entre elles car l'apprentissage et la manipulation de la situation pourraient les faire disparaître, ce qui n'est pas le cas avec les illusions perceptives[199]. En réalité, l'idée d'erreur a déjà été débattue dans le cas de la perception. Comme dans le cas du raisonnement, le courant écologique de psychologie de la perception considère la notion d'erreur et d'illusion comme trompeuse. Encore une fois, nos systèmes perceptifs sont le fruit d'une sélection naturelle qui a abouti à des systèmes fonctionnels dans des contextes écologiques donnés. Comme pour le raisonnement, les psychologues imposent des normes sur notre fonctionnement qui ne sont pas celles ayant garanti la survie de nos ancêtres. Par exemple, percevoir « correctement » les nuances de couleur d'une baie au milieu du feuillage est moins fonctionnel que de percevoir la baie comme ayant la même couleur indépendamment de l'heure du jour –(même si les changements dans le medium (l'air) font que, physiquement, les couleurs sont différentes le matin et le soir). Notre système haptique répond à des quantités physiques qui ne sont pas nécessairement celles capturées par les termes modernes de poids ou de masse, mais plutôt des qualités « mixtes » qui prennent en compte à la fois la masse et le volume d'un objet. Ce qu'un psychologue « cognitiviste » diagnostique comme une erreur de perception – une illusion de couleur ou une illusion de poids – est en réalité, pour un psychologue de la tradition écologique, un signe d'adaptation qui a bien fonctionné : le système perceptif répond à des objectifs et à des qualités physiques fonctionnelles à la survie de l'organisme. [Pour une comparaison des positions « cognitivistes » et « écologiques » aux illusions perceptives, voir : Pasquinelli 2006 (thèse doctorale) et Pasquinelli 2012.]

3.4.3.3 À l'origine de nos erreurs

Le bilan de l'approche « évo-éco » est le suivant : les individus raisonnent bien quand les circonstances sont favorables (au sens où elles sont cohérentes avec les conditions ayant permis un avantage sélectif au module en question). Cette rationalité adaptative (par opposition à une certaine vision de « caractère irrationnel » de notre raisonnement) doit nous motiver pour comprendre l'origine de nos erreurs plutôt que simplement supposer notre irrationalité permanente[200]. Dans ce qui suit, nous regarderons donc plus en détail comment la théorie évolutionniste explique l'existence de biais, et distingue plusieurs types de biais selon leur origine en termes adaptatifs[201].

En effet, même dans le cadre d'une approche « évo-éco », le concept de biais reste légitime. Il est utilisé pour indiquer les déviations d'une norme attendue – qui correspond à la réalité objective (Haselton, Nettle & Murray 2015[202] ; voir aussi, dans un autre domaine, Lea & Ryan 2015). Comme le concept d'illusion, celui de biais est d'ailleurs utile au moins à deux niveaux : il permet d'identifier des conditions (contexte, stimuli) où notre fonctionnement cognitif prend systématiquement une direction « surprenante » ; il permet de mieux comprendre certains aspects structurels de notre fonctionnement (à quel genre de stimuli ou classes de stimuli nous répondons dans certains contextes)[203] (voir : Pasquinelli 2006 (thèse doctorale) et Pasquinelli 2012 pour une discussion concernant l'utilité du concept d'illusion, en dépit des critiques du courant écologique).

Haselton, Nettle & Murray (2015) ont ainsi identifié trois raisons possibles en vertu desquelles des biais peuvent émerger et avoir l'aspect de « défauts de design »[204] :

- la solution fonctionne bien dans la plupart des circonstances, mais elle n'est pas optimale, juste satisfaisante en raison de nos limites ;
- l'erreur est en réalité du côté du psychologue expérimentateur, qui mesure la réponse à un certain type de tâches, alors que la fonction mesurée répond en nature à d'autres buts ; et pour ces buts, elle est précise et efficace ;

- l'erreur a un coût très bas en nature par rapport au coût d'une réponse sans erreur, il peut donc passer le filtre de la sélection naturelle.

L'erreur de type 1 est la plus discutée dans la littérature sur les biais. Qu'il s'agisse d'une solution « *quick and dirty* » ou « *fast and frugal* », une heuristique a des limitations. D'ailleurs, ce genre de solution rapide et à bas coût existe justement parce que nous avons des limitations de calcul (*processing limitations*)[205].

L'erreur de type 2 est l'erreur qui consiste à oublier de considérer le fait que l'optimalité écologique peut avoir un sens différent de celui établi par l'expérimentateur (optimalité écologique, liée au contexte, vs. norme extérieure, logique). Cette catégorie inclut également des situations dans lesquelles des réponses évolutivement anciennes sont déclenchées non à propos par des stimuli récents (évolutivement saillant) qui ressemblent aux stimuli naturels à même d'activer la réponse – éventuellement, le stimulus pouvant se présenter comme un superstimulus : un stimulus plus puissant que celui naturel[206].

L'erreur de type 3 est basée sur l'hypothèse selon laquelle les erreurs peuvent donner lieu à des faux positifs ou à des faux négatifs : ce qui compte, ce sont les coûts de ces erreurs. Un système bien adapté laissera plutôt passer des erreurs qui n'influent pas de façon importante sur les aptitudes (*fitness*), alors que les erreurs désastreuses pour les aptitudes seront corrigées (Haselton, Nettle & Murray 2015[207] ; Haselton & Nettle 2006).

3.4.4 Normes externes et normes internes : retour à l'EC

Traiter les notions de biais et d'heuristique dans une perspective évolutionniste permet de considérer la rationalité d'un point de vue adaptatif plutôt que normatif. Cela signifie-t-il que nous devrions par conséquent abandonner l'idée de normes externes ou standards contre lesquels mesurer les performances humaines ?

Cela rendrait très difficile de définir des objectifs pour l'éducation de l'EC, et des critères d'évaluation corrélés. L'éducation (du moins l'éducation formelle) a en effet un but normatif, lié aux évolutions culturelles de notre espèce. C'est une forme d'ingénierie, qui a pour but de modifier les capacités naturelles, en les outillant et en les développant dans des directions cohérentes avec la culture.

La naturalisation des biais, la notion de rationalité adaptative et, plus généralement, l'approche « évo-éco » ne sont pas contradictoire avec le fait d'établir des normes culturelles de désidérabilité ou de donner une valeur supérieure au critère de respect de la vérité et d'adhérence aux faits. Ces normes constituent des objectifs à atteindre et permettent d'évaluer si les objectifs sont atteints (Haselton et al 2009[208]).
Cependant, si les critères sont établis de façon « externe », à quoi bon mobiliser une approche adaptative et évolutionniste ? Ne pourrions-nous pas nous limiter à éduquer l'EC sur la base d'un certain nombre de normes établies, d'objectifs normatifs, et ignorer les aspects que nous avons présentés jusqu'ici ?

Comme nous l'avons vu, contrairement à l'approche H & B, l'approche « évo-éco » ou de la rationalité adaptative a comme objectif d'expliquer pourquoi il est si difficile de se conformer à certains de ces normes et standards culturellement établis (écart entre normes externes et normes internes, ou erreurs de type 2). L'approche évolutionniste reconnaît en outre que les solutions intuitives qui répondent aux normes internes peuvent avoir des limites et mener à l'erreur (type 1). Les erreurs plus à même d'avoir passé le filtre de la sélection sont celles qui ne comportent pas un enjeu vital mais qui sont à la limite des erreurs de super-prudence (type 3).

Bien que les motivations puissent paraître d'intérêt purement théorique, adopter une approche « évo-éco » entraîne des conséquences directes pour l'éducation de l'EC et également pour notre poursuite de sa naturalisation.
La théorie de l'évolution donne un cadre de pensée pour analyser les biais et les erreurs répertoriés par la psychologie du raisonnement notamment *via* des dispositifs expérimentaux et utilisés pour expliquer des comportements en situation écologique. En comprenant mieux les causes des choix opérés par l'apprenant (en distinguant notamment les causes proximales ou distales : l'influence du contexte, de la manière dont est posé le problème, la finalité implicite poursuivie par l'apprenant), nous pouvons mieux évaluer les limites de son raisonnement et envisager des marges de progression. Nous évitons ainsi certains écueils qui déboucheraient sur des conseils ou des pistes d'éducation inappropriées.

L'approche « évo-éco » montre notamment que nous ne sommes pas dans un combat contre les biais et les erreurs de notre cognition. Il s'agit plutôt de chercher à identifier les capacités cognitives qui nous permettent naturellement de procéder à une évaluation de la qualité épistémique de l'information et de prendre des décisions conformes à la réalité (ou plutôt des décisions optimales, selon la définition donnée de l'optimalité). Ces capacités ne sont pas sans limites. L'approche adaptative nous permet même de prévoir les contextes qui devraient les mettre en difficulté : ceux qui s'éloignent de notre évolution naturelle comme les nouveaux contextes d'information, ou des besoins liés à notre culture. Nous pouvons ensuite chercher à améliorer chacune des capacités identifiées à l'aide d'outils que nous convoquons quand les situations l'exigent.

3.4.4.1 « Ne pas jeter le bébé avec l'eau du bain » : intérêt d'une approche H & B dans le cadre d'une approche cognitive de l'EC

Opposer de façon trop nette les deux approches « *quick and dirty* » and « *fast and frugal* » serait cependant excessif. La critique évolutionniste aux approches de type H & B permet plutôt de mettre en évidence la nécessité de questionner l'origine des erreurs et biais qui affectent la prise de décision et la formation d'opinions et de croyances, afin de cibler des stratégies d'amélioration réalistes plutôt que de simplement admettre ces biais et chercher à les contrer.

Ce type de critique n'élimine pas le fait que, dans certains contextes (type de question, type d'exemple), nous commettons des « erreurs » (voir les réponses de Gilovich & Griffin 2002 et Kahneman & Tversky 1996). Gigerenzer admet lui-même que de vraies illusions cognitives existent également et que le type de « traitement » que nous pouvons appliquer à celles plus artificielles les réduit sans nécessairement les faire disparaître : quelque chose doit donc encore être expliqué, c'est-à-dire comment notre fonctionnement cognitif est induit à donner une réponse qu'il n'aurait pas donné en d'autres circonstances (Gigerenzer, Hoffrage & Goldstein 2008). L'opposition entre les deux approches « *quick and dirty* » and « *fast and frugal* » semble donc excessive à la lumière des faits, car nous avons l'impression que ce qui est surtout critiqué est le fait que le focus de l'attention porte sur un aspect plutôt que sur l'autre des mêmes phénomènes.

3.4.4.2 Que faire avec nos biais et nos erreurs dans le cadre d'une approche éducative de l'EC ?

Deux processus sont souvent indiqués dans la littérature sur le raisonnement comme fondamentaux pour prendre une décision correcte, sans biais :

- la réflexion, lente, coûteuse (opposée aux processus automatiques, rapides, économes de la pensée) ;

- l'éventuelle inhibition, ou éventuel contrôle, sur les réponses basées sur les processus plus automatiques de la pensée.

Une vaste littérature en psychologie du raisonnement décrit le fonctionnement cognitif comme basé sur deux types des processus : des processus de type 1 (T1), rapides et sans effort, et des processus de type 2 (T2), lents et demandant un effort en termes de ressources cognitives, par ailleurs limitées. Vitesse et effort sont les caractéristiques associées avec les deux typologies de processus, mais souvent T1 et T2 sont aussi opposés sur la base de leur relation à l'intentionnalité, la contrôlabilité, l'accès à la conscience et l'efficience (Pennycook et al. 2018).

Dans les années 1970-1980, les théories duales de la pensée (*dual process theories*, DP) alignent de longues listes de caractéristiques toujours associées entre elles (si un processus est volontaire, il est aussi conscient, et ainsi de suite). Les théories actuelles de type DP ont abandonné ces longues listes de caractéristiques. Evans et Stanovich, par exemple, qui appartiennent aux représentants les plus influents des théories DP, séparent les caractéristiques définitoires de T1 et T2 des caractéristiques pouvant y être potentiellement associées (Evans & Stanovich 2013). Ils identifient comme caractéristiques définitoires des deux types de processus : le fait qu'ils demandent ou non de la mémoire de travail et le fait qu'ils soient ou non autonomes – qu'ils demandent ou non un contrôle attentionnel. Les auteurs reconnaissent aussi que les listes de caractéristiques des théories DP traditionnelles sont souvent incohérentes : par exemple, les processus de type 1 ne sont pas nécessairement plus anciens évolutivement que les autres ; les processus de type 1 peuvent amener à des réponses efficaces et efficientes, ceux de type 2 à des biais et des erreurs. Comment s'articulent entre eux les processus de type 1 et 2 ? Selon l'une des approches duales, l'approche « défaut-interventionniste », lorsque nous sommes confrontés à un problème nouveau, les réponses intuitives s'activent automatiquement et sans effort. Si nous ne sommes pas préparés pour ce type de situation, les réponses peuvent alors être inappropriées. Il faut donc intervenir de façon réflexive pour dépasser la réponse intuitive. Or, souvent, nous faisons preuve d'une certaine paresse : nous choisissons des indices faciles à identifier et à reconnaître et ne cherchons pas les caractéristiques plus précises de la situation. Ceci nous amène à choisir la mauvaise réponse (Evans & Stanovich, 2013).

Plusieurs critiques peuvent être émises envers les modèles dualistes, et notamment envers ceux de type défaut-interventionniste (Mercier & Sperber 2011 ; Pennycook et al. 2018 ; Bago & De Neys 2019[209]). Des données expérimentales récentes jettent un doute sur les deux hypothèses de la cécité au biais et de la nature uniquement corrective des processus « algorithmiques » (T2). Plusieurs études ont par exemple mis en évidence que, même quand ils tombent dans le « biais » lors de tâches classiques de logique, les sujets peuvent détecter l'existence d'un conflit, et donc être sensibles à l'erreur (Bago & De Neys 2019[210] ; De Neys, Lubin & Houdé 2014[211]). En effet, la confiance (mesurée *via* des échelles verbales subjectives et explicites, proposées après que la réponse a été donnée) est plus faible dans le cas de tâches avec conflit que dans des tâches sans conflit, et ceci même pour les sujets donnant la mauvaise réponse. D'autres résultats montrent qu'une majorité des sujets qui donnent la bonne réponse après réflexion (condition lente) ont déjà donné la bonne réponse dans la condition rapide. Dans le cas de ces participants, l'intuition initiale est (et reste) correcte. Ces données invitent à repenser le rapport entre T1 et T2 : T1 pourrait donner des réponses intuitives correctes même d'un point de vue logique ; T2 permettrait plutôt de justifier *a posteriori* et de communiquer son intuition (Bago & De Neys, 2019[212]).

Par conséquent, des modèles plutôt hybrides que duaux de la pensée ont alors récemment émergé (Bago & De Neys 2017, 2019, mais aussi : Ball, Thompson & Stupple 2017 ; Białek & De Neys 2017 ; Pennycook 2017 ; Pennycook, Fugelsang, & Koehler 2015 ; Newman & Thompson 2017 ; Trippas & Handley 2018). Dans ces modèles, une réponse « rationnelle » peut émerger dès le départ, au niveau des processus de type 1. Le choix entre des réponses intuitives concurrentes dépend de la force avec laquelle elles sont respectivement activées

(Bago & De Neys 2019). Puisque les deux réponses sont simultanément « présentes à l'esprit », un conflit est détecté entre les deux. Le sujet expérimente une sensation de confiance envers la réponse donnée d'autant plus forte que l'autre réponse est moins activée. Cette hypothèse permet de rendre compte des différences interindividuelles : différents sujets peuvent avoir des intuitions logiques plus ou moins fortes, selon la facilité avec laquelle leur réponse logique est activée. Cette hypothèse rejoint en outre la théorie argumentative du raisonnement proposée par Mercier & Sperber (2011, 2016) et par Trouche, Sander & Mercier (2014). Le rôle du fonctionnement de type de 2 serait en effet celui de permettre de trouver une justificaiton à ses choix de réponse, et de la communiquer à autrui, éventuellement pour convaincre.

Une autre question qui se pose dans le cadre des modèles duaux (y compris hybrides) de la pensée consiste à savoir comment nous passons d'un fonctionnement d'un type à un autre (en réalité, du fonctionnement de type 1 au fonctionnement de type 2), et qu'est ce qui détermine la stratégie gagnante (celle qui va déterminer la décision).

Une réponse classique à la question de la mise sous contrôle des processus plus automatiques et réflexifs de la pensée, en faveur de ceux plus réflexifs, fait entrer en jeu les mécanismes d'inhibition. Plutôt que de parler de deux systèmes, Stanovich propose alors un modèle en trois typologies de processus : des processus de type 1, rapides ; des processus de type 2, algorithmiques et qui peuvent se substituer à ceux de type 1 ; des processus de type 3, réflexifs, qui contrôlent la mise en place des processus de type 2 et inhibent ceux de type 1. Ce troisième type de processus est plus une disposition qu'une capacité : il s'agit d'une disposition à chercher plus d'informations, à prendre différents point de vue, etc. et à permettre ainsi aux capacités cognitives algorithmiques de se mettre en place (Evans & Stanovich, 2013).

3.4.4.3 Inhiber ?

Le concept d'inhibition permet de penser la relation entre les différentes stratégies coexistant dans l'organisme pour la résolution d'un problème, mais il est également problématique (Aron 2007).

Dans son acception primaire, l'inhibition est un processus neurophysiologique impliqué dans le contrôle central de reflexes moteurs et la propriété de neurones qui communiquent *via* des classes particulières de neurotransmetteurs, les neurotransmetteurs inhibiteurs (par exemple, GABA) dont l'action inhibe celle du neurone cible. L'effet de ces processus cellulaires est donc celui d'une modulation de l'activité neurale. Le concept d'inhibition est également présent en psychologie, avec des significations qui ne coïncident pas nécessairement avec celles neurophysiologiques (Aron 2007). L'inhibition y est notamment traitée comme une fonction de type exécutif, parfois appelée « contrôle cognitif ». Les fonctions exécutives sont définies comme des fonctions d'ordre supérieur, chargées de contrôler la mise en place de fonctions de plus bas niveau. Il n'existe pas une classification unique des fonctions exécutives, qui peuvent comprendre plusieurs des fonctions suivantes : contrôle exécutif, mémoire de travail, contrôle attentionnel, flexibilité cognitive, capacité de planification. Une classification influente les réduit néanmoins à trois : contrôle inhibiteur, mémoire de travail, flexibilité cognitive (voir Diamond & Ling 2016[213]).

Le contrôle exécutif peut être mis en évidence de manière expérimentale par des tâches particulières, notamment par des tâches d'amorçage négatif. Imaginons qu'un sujet doive réagir à un objet O1 – un lapin – et ignorer un objet O2 également présent – un chat. Dans un second essai, qui suit immédiatement le premier, le sujet doit répondre à un nouvel objet O3 – un chien – ou alors à l'objet précédemment bloqué – O2, le chat. Dans ce dernier cas (O2), mais non dans le cas du nouvel objet (03), la réponse est plus lente. Cette caractéristique de la réponse est interprétée comme un effet résiduel d'un mécanisme préalable d'inhibition de la

réponse 02 (Neill, Valdes & Terry 1995 ; Tipper 2001). Ici, le contrôle exécutif agit sur l'attention, notamment sur l'attention sélective. De façon plus générale, le rôle du contrôle inhibiteur ou cognitif serait de permettre de prendre des décisions (y compris des décisions au niveau subpersonnel, perceptives ou motrices par exemple) sans être complètement dominé par les stimuli externes qui arrivent en continu et qui peuvent automatiquement déclencher des actions en réponse (Tipper 2001). Les patients présentant des lésions du cortex préfrontal montrent par exemple un comportement pouvant s'interpréter comme un manque de contrôle inhibiteur, car ils tendent à prendre en main tout objet qui est posé devant eux (Lhermitte 1983). Un siège important du contrôle exécutif serait donc dans le cortex préfrontal (PFC). Des études en neuro-imagerie ont mis en évidence le rôle spécifique du cortex frontal inférieur, notamment du gyrus inférieur (IFG), dans le raisonnement de type logique et dans la perturbation de celui-ci (Aron, Robbins & Poldrack 2014). Parallèlement aux études sur le raisonnement et ses soubassements neuraux, Adele Diamond et d'autres psychologues développementaux ont utilisé le concept d'inhibition – de manque d'inhibition – pour expliquer certains comportements de l'enfant comme les réponses au test piagétien A – non-B (l'enfant cherche et trouve plusieurs fois un objet caché à l'endroit A. Quand l'objet est déplacé devant ses yeux dans le lieu B, l'enfant continue à chercher en A) (Diamond 1991).
Il faut néanmoins savoir que le concept d'inhibition ne va pas de soi dans la littérature en psychologie et qu'il fait l'objet de plusieurs critiques.

Premièrement, tel qu'il est présenté, notamment dans le cadre des théories duales de la pensée, il implique l'existence d'une sorte d'entité supérieure au sein du fonctionnement cérébral qui déclencherait l'inhibition afin de mettre sous contrôle la réponse automatique et laisser libre cours à celle algorithmique. Le modèle comportant trois systèmes - proposé par Stanovich - ne fait que repousser le problème du déclenchement de l'inhibition.

Deuxièmement, l'émergence de réponses logiquement correctes peut également être expliquée par des mécanismes autres que l'inhibition : par exemple, dans le cadre des théories hybrides discutées ci-dessus, par des mécanismes de renforcement des réponses désirables ou de l'information pertinente. Ce renforcement permettrait aux réponses correctes ou désirables de passer le seuil de la décision et donc d'être choisies, même si d'autres – normalement plus automatiques – existent (Bago & De Neys 2019).
Troisièmement, même les effets des lésions du PFC peuvent être expliquées par d'autres mécanismes que l'inhibition, par exemple en postulant que ces lésions comportent des perturbations des fonctions permettant le bon déroulement de la tâche, plutôt que des perturbations d'un système d'inhibition (Aron 2007). Certains auteurs considèrent donc que le concept d'inhibition est plutôt un raccourci utilisé pour décrire des situations où deux ou plusieurs réponses possibles sont en conflit. Ce conflit est alors détecté et un choix est effectué, avec implication de la mémoire de travail ; les mécanismes permettant ce choix et le déroulement d'une bonne performance ne relèvent pas nécessairement de l'inhibition de l'une des réponses ou des stimuli impliqués. L'utilisation du terme « inhibition » dans toutes ces situations pourrait amener à confusion car, par exemple, il cache l'existence de mécanismes modulateurs *top-down* qui agissent plus comme des amplificateurs que comme des inhibiteurs du signal. L'inhibition devrait donc plutôt être considérée comme l'un des mécanismes en jeu dans ce type de modulation et de contrôle de la réponse, et le concept d'inhibition être évoqué seulement dans les circonstances appropriées. Parmi celles-ci, le contrôle moteur, mais aussi le contrôle sur les émotions, attention et mémoire semblent porter les signes de processus inhibiteurs actifs, qui bloquent des réponses déjà initiées au niveau central (Aron 2007). La neuro-imagerie semble être une voie particulièrement adaptée pour distinguer différents mécanismes potentiellement en jeu dans la modulation des réponses observées au niveau comportemental (Aron 2007).

Quatrièmement, les efforts pour « entraîner » une éventuelle capacité générale d'inhibition à travers des activités « classiques » - comme la tâche de Stroop ou la tâche *Go-no-Go* - ne donnent pas de résultats univoques, notamment lorsque nous nous intéressons au transfert de l'entraînement sur des tâches plus ou moins éloignées (Enge et al. 2014[214] ; voir aussi : Simons

et al. 2016). Ainsi, dans une méta-analyse de 19 études concernant le contrôle inhibitoire (*via* entraînement par tâches classiques comme *Go-no-Go*), Allom, Mullan & Hagger (2016[215]) trouvent un petit effet positif avec transfert sur des comportements liés à la santé, mais sur des tests de transfert proposés immédiatement après l'entraînement. Diamond & Ling (2016) ont dédié une revue de littérature aux différentes interventions concernant les fonctions exécutives, et décelé des régularités par rapport aux interventions qui se montrent efficaces. La première considération concerne le fait que les fonctions exécutives peuvent être améliorées, et que les interventions efficaces se situent autant du côté des interventions de type cognitif que des interventions qui comportent de l'activité physique (arts martiaux par exemple). D'autres considérations concernent le type d'effets et la manière de mener les activités efficaces. Le transfert est étroit : entraîner une fonction exécutive comme la mémoire de travail ne donne pas des avantages en termes de contrôle inhibitoire ou de flexibilité cognitive. Gagnent plus ceux qui partent de performances faibles. Le temps a également un rôle important : une pratique prolongée est toujours plus efficace qu'une pratique de courte durée ; de plus, quand la pratique s'arrête, les effets peuvent disparaître. L'effort joue aussi un rôle : continuer à pratiquer une fois que la tâche est devenue facile n'a pas d'apports, la tâche devant rester un défi pour avoir un effet sur les fonctions exécutives (comme pour le gain d'expertise en général). Une activité censée entraîner les fonctions exécutives qui s'est montrée efficace dans un certain contexte n'est pas nécessairement toujours efficace, son efficacité dépendant en effet de la manière dont elle est menée et pratiquée : la même intervention peut être efficace dans un contexte mais pas dans un autre. L'exercice physique irréfléchi ne se montre pas efficace. En revanche, l'exercice physique accompagné de réflexion, planification, résolution de problèmes montre des effets positifs. Quant aux entraînements cognitifs, le plus connu concerne la mémoire de travail (CogMed). Il a montré des effets positifs (mais pas de transfert à distance et sur des tâches de la vie réelle) mais uniquement en présence d'un moniteur qui accompagne l'entraînement à l'ordinateur, en le rendant ainsi réflexif. Ce qui amène à douter ultérieurement de l'efficacité des entraînements dont la caractéristique est d'être implicites, non réflexifs[216].

Houdé & Borst (2015) étudient – *via* des méthodes de neuro-imagerie – le rôle des processus inhibiteurs dans le raisonnement logique (règle *si-alors* dans des tâches perceptives) et montrent des activations compatibles avec l'hypothèse de l'inhibition.
Houdé & Borst (2015) montrent en outre que l'instruction qui implique un *feedback* négatif sur la mauvaise réponse est plus efficace – favorise l'apparition de la bonne réponse – que la simple instruction positive qui vise à renforcer la « bonne réponse ». Les auteurs appellent ce *feedback* négatif « inhibition »[217]. Cependant, le *feedback* est explicite et les auteurs considèrent donc que la métacognition explicite (qui favorise la mise en place du processus d'inhibition) joue un rôle dans le succès de l'instruction.
En conclusion, nous pouvons légitimement considérer le concept d'inhibition comme le seul processus, ou fonction, impliqué dans la modulation des comportements, notamment en présence de choix à effectuer entre réponses, et comme un concept unique. Nous pouvons également envisager l'efficacité de l'entraînement implicite de l'inhibition comme une méthode pour réduire les erreurs dans le raisonnement à différents niveaux et dans plusieurs contextes, sauf si celui-ci prend la forme d'un enseignement métacognitif explicite. Néanmoins, à la lumière des connaissances présentes, le concept restreint d'inhibition et les études sur les mécanismes spécifiques de contrôle inhibitoire restent intéressants pour une théorie opérationnelle de l'EC, notamment une fois couplés avec une approche métacognitive explicite.

3.4.4.4 Prendre conscience des biais les plus communs et des circonstances dans lesquelles ils peuvent nous mettre en difficulté

Nous avons établi que les biais ne sont pas en soi des erreurs à éliminer de façon systématique. Un entraînement à l'inhibition qui rendrait les élèves globalement moins susceptibles de suivre leurs intuitions ou automatismes ne serait donc pas souhaitable. Nous

venons de voir que ce type d'entraînement, pensé comme une forme implicite de gymnastique du contrôle inhibitoire, n'est pas non plus efficace. Deux questions restent ouvertes. La première concerne l'utilité pragmatique d'une prise de conscience (explicite, métacognitive) des limites de notre propre fonctionnement cognitif, là où les erreurs sont plus fréquentes ou dommageables. La deuxième analyse les effets sur la capacité d'évaluer correctement la qualité épistémique des informations à notre disposition, et de calibrer notre confiance en conséquence.

Prenons le cas concret d'une réaction commune, qui consiste à interpréter – peut-être trop rapidement – une relation de corrélation comme impliquant une connexion causale. L'évaluation correcte de la qualité d'une information disponible nécessite de s'enquérir de la manière dont la relation causale est établie, par des méthodes aptes à diminuer l'incertitude ou non. Cependant, prendre conscience que nous pouvons facilement être induits à utiliser le seul indice de corrélation comme entraînant la causalité pourrait nous aider à comprendre pourquoi des méthodes appropriées sont nécessaires, de quelle manière elles diminuent les chances d'erreur dans un jugement de causalité, et donc à nous motiver pour y avoir recours ou pour les prendre en compte dans nos jugements de qualité de l'information. Est-ce le cas ?

La littérature concernant la prise de décision éclairée est peu optimiste, encore une fois, quant aux effets des entraînements (Marewedge et al. 2015[218]). Elle tend à montrer que le transfert d'un domaine à un autre représente un vrai défi (nous reviendrons sur cette problématique lorsque nous discuterons de l'éducabilité de l'EC). *Quid* des entraînements appelés de « débiaisage » (*debiasing*) (Fischoff 1982) ? Ces derniers consistent le plus souvent en des exercices permettant de comprendre qu'un certain type d'information tend à ne pas être pris en considération lorsque nous émettons des jugements. Le cas plus célèbre de succès est celui des activités de débiaisage proposées par Richard Nisbett et ses collègues dans le cas d'enseignements de statistiques, probabilités et biais cognitifs pour des élèves de médecine et de psychologie (Nisbett et al. 1987). Ces interventions ont en effet été évaluées à l'aide de tests, suivant des intervalles de temps et en dehors des conditions de laboratoire, et ont montré des effets de transfert sur des tâches qui demandent l'application d'outils statistiques et logiques appris en cours. Il existe aussi des cas de succès avec une seule intervention (vidéo illustrant des biais et activité de simulation). Le succès est mesuré immédiatement après l'intervention et à distance mais *via* des tâches de transfert proche (Marewedge et al. 2015). Aczel et al. (2015) ont mis en place un entraînement en plusieurs phases : l'apprenant est d'abord confronté à ses erreurs *via* un test et un *feedback* spécifiques. Puis, il apprend à reconnaître les similitudes entre différentes situations, et à en extraire un principe commun : un biais. Il est ensuite poussé à chercher l'équivalent dans ses mémoires autobiographiques et instruit à propos de la nature du biais. À ce moment-là, il apprend aussi des stratégies pour y répondre et doit se projeter dans le futur pour trouver des situations dans lesquelles il va réutiliser cet apprentissage. Le post-test a lieu un mois après l'entraînement et utilise les mêmes questions sur lesquelles les sujets ont été testés au début pour découvrir leurs biais. Aczel et al. (2015[219]) analysent aussi différentes catégories de techniques de débiaisage et mettent en évidence les difficultés généralement liées au transfert. Enseigner l'existence d'un biais qui affecte la décision, notamment lorsqu'il reste à un niveau abstrait, ne permet pas nécessairement à l'apprenant de reconnaître les situations concrètes dans lesquelles le biais se présente et de savoir comment réagir. L'apprentissage avec des exemples et dans des domaines d'application spécifiques est lié à l'apprentissage de ces contextes, notamment dans le temps. La similarité de surface entre situations d'apprentissage et situations d'application facilite le transfert, mais cette condition met de côté beaucoup de situations ou contenus non envisagés. En outre, apprendre la nature des biais sans développer les stratégies pour les dépasser (par exemple, des stratégies liées à l'application de statistiques plus sophistiquées que celles employées spontanément) ne permet pas de réagir de manière appropriée, même si le biais est reconnu. Nous discuterons dans la partie dédiée à l'éducabilité de l'EC des possibles solutions pour faciliter le transfert, mais les objections soulevées ici suggèrent que le débiaisage n'est efficace que dans des conditions très particulières.

3.4.5 Heuristiques, erreurs récurrentes et biais. Des indications pour l'éducation de l'EC

L'EC consiste – selon la définition que nous en avons donnée – à évaluer le plus correctement possible la qualité d'une information et à se former un jugement – implicite ou explicite – de confiance calibré par rapport à cette information, en vue d'une prise de décision.

Cette définition est volontairement restreinte. Cependant, elle laisse de la place pour la recherche de composantes cognitives spécifiques ou *building blocks*. Nous avons préalablement identifié deux composantes cognitives de l'EC : les processus et mécanismes de la vigilance épistémique et ceux de l'épistémologie naïve et de la sensibilité métacognitive. Cette dernière consiste notamment à évaluer la confiance dans une information ou décision sur la base des preuves disponibles. Cette évaluation constante peut générer des sentiments de confiance au niveau personnel, et informer le jugement. Mais cela peut aussi se traduire uniquement en une caractéristique des activations neurales par rapport à certains stimuli, et donc seulement agir à un niveau subpersonnel, fonctionnel. La discussion concernant la notion d'erreur et de biais montre encore une fois l'importance – en évaluant la qualité des informations – de prendre en considération son fonctionnement cognitif, et pas uniquement le sien.

De notre discussion concernant les heuristiques et biais émerge en effet la considération selon laquelle dans la cognition humaine opèrent des attracteurs universels qui rendent certaines informations plus séduisantes, certaines idées plus convaincantes, ou plus compréhensibles, attirantes, saillantes, mémorables – et ceci indépendamment de leur valeur de vérité. Prendre conscience de ces attracteurs, mieux connaître le fonctionnement cognitif (pas uniquement le sien, mais le fonctionnement de la cognition en général) permettrait de mieux anticiper les chances de commettre une erreur de jugement par rapport à l'information disponible et à sa qualité.
Est-ce que cela signifie que l'EC consiste à apprendre à dompter tous ses biais, à résister, à inhiber ses réponses spontanées et rapides, c'est-à-dire ses intuitions ? Comme nous l'avons dit plus haut, ce que nous appelons couramment des « biais », avec une intention négative, conduisent à des erreurs seulement dans certains contextes et circonstances. Une inhibition générale des biais ou du fonctionnement du soi-disant « système 1 de la pensée » n'est donc pas souhaitable en soi.

Cependant, s'il est pris dans une acception large, le concept d'inhibition peut aussi inclure des stratégies et des connaissances qui ne bloquent pas directement une réponse indésirable, mais qui augmentent les chances de faire surgir une réponse désirable. Ces stratégies et connaissances doivent avant tout être présentes à l'esprit, donc avoir été apprises si nous voulons espérer les utiliser à un moment donné. De plus, leur accès doit être facilité, donc elles doivent être automatisées ou presque dans leur utilisation, afin d'être utilisées au plus bas coût possible. Nous avons vu l'exemple des connaissances en statistiques, qui constituent des outils permettant de contrecarrer des erreurs récurrentes dans notre appréciation spontanée du rôle du hasard, du risque lié à des événements.

Reste à s'interroger sur l'importance de savoir reconnaître différents types de biais pour les anticiper et en réduire ainsi les effets négatifs éventuels sur l'évaluation de la qualité des informations disponibles et du calibrage de la confiance en vue d'une prise de décision. Si la littérature nous donne une vision sceptique quant aux chances (et à l'opportunité) de débiaiser de façon générale la cognition humaine, elle laisse la place pour des interventions pédagogiques permettant de prendre conscience des biais et d'utiliser sa métacognition pour mieux identifier les situations où le risque de commettre des erreurs d'évaluation est plus fort.

Un exemple servira à mieux éclaircir ce point de vue. Une vaste littérature est dédiée au biais de confirmation : la tendance à rechercher des informations conformes à nos opinions et à les évaluer de manière plus bégnine que nous ne le ferions avec des informations de signe contraire, plus une résistance à changer nos opinions pour embrasser celles des autres. Nous avons décrit l'incertitude – donc le manque de confiance – comme un moteur permettant de changer de position et de dépasser le biais de confirmation. Cependant, connaître l'existence et la force de ce biais permet également de se forcer la main, de façon volontaire. Par exemple, en créant des communautés de paires qui évaluent et critiquent nos arguments à notre place. C'est le cas de la création de groupes de consensus, de comités scientifiques de paires, mais aussi de groupes de travail coopératif. Connaître les mécanismes sous-jacents du biais de confirmation et d'autres biais de nature sociale permet d'anticiper que les groupes qui fonctionneront mieux du point de vue du dépassement du biais de confirmation sont des groupes hétérogènes, où existent des opinions contraires. Cette prévision est d'ailleurs confirmée par certaines études sur le raisonnement de groupe (par exemple, Mercier 2016), même s'il faut lier ces connaissances à d'autres concernant les effets de prestige, de conformité, etc.

De ces considérations émergent donc des indications pratiques pour l'éducation de l'EC. La principale consiste à outiller les capacités d'évaluation de la qualité de l'information et celles de calibrage de la confiance grâce à des connaissances appropriées, capables de contrecarrer nos penchants naturels, lorsque ceux-ci ont des chances importantes de nous mener en erreur.
Il s'agit donc :

- d'apprendre (de faire apprendre) à reconnaître les situations dans lesquelles, en raison de notre fonctionnement cognitif naturel, nous sommes plus à même de nous tromper – dans le sens de donner trop ou trop peu de poids à une information dans le cadre d'une prise de décision, de ne pas juger l'information sur la base de sa qualité mais d'influences de différentes natures autres que la qualité objective de la preuve et de la source ;
- d'apprendre à utiliser de manière fluide, avancée, des stratégies plus puissantes ou plus adaptées que celles naturelles comme : les stratégies qui emploient des outils statistiques, des probabilités ; le recours à des outils de coopération sociale afin de mettre sous contrôle, de façon volontaire, des biais comme celui de confirmation.

3.5 Synthèse et remarques générales

Les mécanismes de la vigilance épistémique, sensibilité métacognitive, prise de conscience et contrôle sur les biais les plus communs constituent autant de bases cognitives naturelles de l'EC, en tant que **capacité à évaluer la qualité épistémique des informations disponibles et de calibrer sa confiance dans ces informations, en vue d'une prise de décision**. **Ces mécanismes peuvent être mis en échec dans des situations complexes, nouvelles (du point de vue de l'évolution), ou tout simplement en vertu de leurs limites intrinsèques.**

L'effort de naturalisation de l'EC ne doit pas être vu uniquement sur un plan théorique. Se baser sur une meilleure connaissance des processus cognitifs est fondamental dans la création d'interventions pédagogiques fondées sur les preuves, l'étape suivante étant naturellement celle de la validation par le test expérimental d'abord à petite puis à grande échelle.

La définition fournie de l'EC et l'investigation de ses bases cognitives nous ont ainsi permis de donner des indications pratiques pour l'éducation, spécifiques par rapport à chaque *building block* cognitif identifié. Nous pouvons maintenant avancer des considérations d'ordre général concernant l'éducation de l'EC.

En termes d'ingénierie pédagogique, l'approche proposée ici vise à **remédier aux limites naturelles des capacités d'évaluation des informations (de première et de seconde main) et à en étendre la portée**. Il ne s'agit pas de construire à partir de rien (*from scratch*) un penseur critique idéal, mais de comprendre ce qui – dans le fonctionnement de nos mécanismes de vigilance épistémique – nous met en difficulté dans l'évaluation de l'information en circulation. Nous parlons donc d'un **EC outillé**, par rapport à un **EC naturel**, qui ferait partie de notre bagage cognitif. Outiller l'EC naturel permet d'accéder à des niveaux de plus en plus avancés d'EC. L'EC expert comporte des outils spécifiques à des domaines d'expertise : par exemple, une personne dotée d'EC expert en médecine est capable d'évaluer des informations spécialisées dans ce domaine, de connaître les sources les plus fiables dans ce domaine, etc.

La question éducative est alors : comment outiller l'EC naturel ? **Améliorer l'évaluation de la qualité épistémique des informations passe par la transmission des critères et des connaissances permettant une évaluation plus adaptée aux situations complexes.**

Les outils pour l'EC avancé comprennent :

- des critères avancés d'évaluation des preuves et des sources, des connaissances pour améliorer l'évaluation de la pertinence et de la plausibilité des informations ;
- des stratégies et des connaissances permettant de mieux calibrer la sensibilité métacognitive et de réduire l'impact des biais métacognitifs. Par exemple, l'amélioration des connaissances de domaine, le recours au *feedback* externe, l'explicitation de l'évaluation de la confiance afin de ne pas utiliser que des indices indirects de façon silencieuse ;
- la prise de conscience des influences cognitives qui peuvent nous amener à donner des évaluations moins justes des informations disponibles, de quand ces influences représentent un risque, de la manière de les contrecarrer grâce à des stratégies et à des outils culturels (par exemple, statistiques) ou artificiels (par exemple, la création de situations de coopération).

Nous préconisons donc pour l'éducation de l'EC :

- de s'appuyer sur les capacités existantes en identifiant clairement leurs limites, notamment les conditions dans lesquelles ces capacités deviennent sous-performantes par rapport aux objectifs établis. Il s'agit par exemple de se doter de stratégies pour identifier l'origine d'une information, de critères pour jauger correctement l'expertise tout en tenant compte du fait que certains indices spontanément utilisés peuvent prendre une valeur différente dans de nouveaux contextes, de critères pour comprendre comment la connaissance est produite ;
- d'identifier clairement les connaissances qui permettent d'outiller les capacités naturelles afin de les rendre plus performantes dans les contextes dans lesquels elles peuvent devenir sous-performantes ;
- d'outiller les apprenants à travers ces connaissances, en les aidant à se les approprier ;
- de préparer les apprenants à reconnaître les contextes pouvant les mettre en difficulté et à transférer leurs outils dans ces nouveaux contextes. Puisqu'il est difficile de s'auto-évaluer, les évaluations externes constituent des aides au calibrage de la confiance. Des évaluations en amont (évaluations diagnostiques) peuvent être utiles pour faire prendre conscience à l'apprenant de ses propres limites et ainsi motiver la recherche de meilleurs critères.

Favoriser une pensée plus lente et un fonctionnement plus contrôlé peut paraître un levier évident pour l'éducation de l'EC. Cependant, prendre le temps ou réfléchir plus et chercher plus d'arguments pour justifier une position ne garantit pas toujours de parvenir à identifier

les bons arguments. Inhiber de façon générale ses premières intuitions n'est pas une stratégie à poursuivre, d'autant plus que celles-ci peuvent être correctes. De même, le fait d'obtenir plus d'information peut devenir un exercice infini si nous ne savons pas quand nous arrêter.

L'objectif final reste de permettre au sujet de mieux calibrer sa confiance relativement aux informations à sa disposition et aux décisions qui en découlent. Il ne s'agit pas de devenir généralement plus méfiant ou de tout vouloir vérifier par soi-même.

En opposition à ces modalités éducatives basées sur des indications « généralistes », nous proposons d'outiller l'EC par des critères plus appropriés d'évaluation de l'information et de calibrage de la confiance que nous portons à ces informations. il s'agit également de mesurer le degré de succès des actions éducatives sur l'EC en mesurant l'ajustement du calibrage de la confiance relativement à l'information disponible.
Il faut remarquer que posséder des connaissances de domaine semble être une clé pour mieux évaluer contenus et sources d'information, tout comme pour mieux calibrer sa confiance dans le domaine en question. Cette considération prône une éducation de l'EC qui ne soit pas découplée d'une acquisition de contenus de connaissance riches et profonds. En pratique, il ne s'agirait pas d'enseigner à bien penser, mais d'enseigner de manière à réussir à bien penser, y compris grâce à un bagage de connaissances. Certaines connaissances acquises dans le domaine en question peuvent être immédiatement transférables dans d'autres domaines, au prix d'un effort d'adaptation et de stratégies efficaces d'enseignement. Mais comment favoriser le transfert d'un domaine d'exercice de l'EC à un autre ? C'est à cette question que nous allons chercher à répondre dans la partie dédiée à l'éducabilité de l'EC.

4. Éducabilité de l'esprit critique

4.1 Éduquer des capacités générales

Est-il possible d'apprendre à penser (de façon plus critique) ? Cette question mobilise une idée longuement débattue concernant l'existence même de capacités générales telles que réfléchir, résoudre des problèmes, être créatif… Pouvons-nous parler de capacités générales qui, une fois acquises, seront appliquées à des contextes et des contenus différents et non liés entre eux ? Est-il possible d'apprendre à penser ? N'apprenons-nous pas plutôt à penser des contenus donnés dans un contexte donné ? Et si nous n'apprenons que dans un contexte, savons-nous nous en débarrasser pour transférer la capacité ou le savoir acquis à un autre contexte, plus ou moins éloigné ? Ces questions ont reçu des réponses parfois opposées. Par conséquent, l'opportunité d'enseigner des capacités générales comme celle de penser de façon critique a tour à tour été affirmée, niée, puis réaffirmée, en mettant alternativement l'accent sur les savoirs locaux, à préférer (ou pas) à ceux généraux. Transférer des savoirs et des capacités appris est possible mais comporte des exigences et des contraintes. Loin de constituer un problème, ceci permet au contraire de fournir des indications pratiques sur la manière de concevoir une éducation de l'esprit critique qui se voudrait transférable vers de nouveaux contextes, notamment dans la vie quotidienne.

Parmi les interventions visant le développement de l'esprit critique, celles qui se montrent efficaces – dans le temps et de façon indépendante du contexte initial d'apprentissage – répondent à ces mêmes critères. Cela donne une première réponse à la question si nous pouvons apprendre à penser de façon plus critique : oui, mais des conditions précises doivent être respectées pour que cet apprentissage soit utilisable ultérieurement.

4.2. L'EC est-il une capacité générale ?

Il était une fois un chef d'État d'une contrée lointaine. Le chef sentait que son pays allait bientôt être menacé par des agressions par des nations voisines, bien mieux dotées que la sienne d'un point de vue militaire. Il fallait être plus malin et chercher à gagner la bataille par l'intelligence plutôt que par la force. Heureusement, le pays pouvait compter sur le plus grand joueur d'échecs du monde ! Le chef d'État était satisfait. Voilà la solution : recruter ce magnifique intellectuel, lui apprendre des rudiments de politique et de techniques militaires et, grâce à son génie, vaincre l'ennemi.

Qu'en pensez-vous ? Le chef d'État a-t-il eu une bonne idée, car les capacités (générales) du maître d'échecs se transposeront automatiquement au champ de bataille et plus globalement à la vie réelle ? Cette tentative est-elle vouée au désastre, parce que les capacités acquises en devenant maître d'échec sont locales et contextualisées à des mouvements de pièces sur un échiquier ? L'idée est-elle prometteuse, mais pas assez réfléchie pour aboutir à quelque chose ?

Cette expérience de pensée est placée au tout début d'un article de psychologie de l'éducation devenu désormais un classique sur les questions de transfert des apprentissages, écrits par des professeurs de Harvard dans les années 1980 (Salomon & Perkins 1988). La question des capacités générales et de leur transfert avait déjà une longue histoire quand l'article a été écrit (à la fin des années 1980). Cette histoire croise à la fois le domaine de l'éducation et celui des sciences cognitives. Au début du XXe siècle, la question est posée par Edward Thorndike, père de la psychologie béhavioriste, qui s'intéresse beaucoup à l'éducation. C'est d'ailleurs dans ce

domaine qu'il mène ses recherches sur le transfert, pour comprendre si nous pouvons considérer l'apprentissage du latin et celui des mathématiques comme étant propédeutiques à d'autres apprentissages et capacités. Dans les années 1950, la question devient pressante, notamment avec le développement des programmes de recherche sur l'intelligence artificielle, dont l'objectif initial est de reproduire des systèmes intelligents capables de résoudre des problèmes ou de « raisonner » sur toute sorte de contenus à partir d'un nombre restreint d'heuristiques générales qui s'appliqueraient indépendamment du contexte occasionnel choisi.

4.2.1 Le problème de l'intelligence ne se pose pas

La problématique des capacités générales croise également un autre domaine de recherche et de long débat : celui concernant le quotient d'intelligence et le facteur g ou d'intelligence générale introduit par Spearman en 1904. Cette idée répond à une intuition assez répandue :

« *La notion d'intelligence générale repose sur une intuition assez partagée, selon laquelle on distingue facilement les individus que tout le monde qualifie d'intelligents de ceux qui le sont beaucoup moins. Toutefois, à y regarder de plus près, on constate que les capacités et les talents peuvent être multiples ; celui qui excelle dans le maniement des subtilités du langage peut utiliser moins bien le raisonnement abstrait, alors que tel autre brillant en mathématiques est incapable de gérer sa vie au quotidien. Néanmoins, l'observation de ces différences individuelles n'enlève rien à l'intuition d'une forme d'intelligence générale qui s'appliquerait à de nombreux domaines de la vie. De fait, les données recueillies depuis des siècles sur les tests d'intelligence – un test complet comporte en général plusieurs sous-tests individuels – confortent cette intuition.* » (Ramus, 2012, p. 4).

Le grand public, et celui des enseignants plus particulièrement, connaît bien une critique portée par Howard Gardner, de l'université de Harvard, à l'idée d'une intelligence générale et d'un score, le QI, qui la quantifie. Selon Gardner, les tests de QI ne mesurent pas des capacités comme celles sociales qui jouent pourtant un rôle important dans la vie quotidienne. Tout en étant correcte, cette objection n'enlève rien au fait que la mesure du QI est fiable (les performances à tous les tests et sous-tests sont corrélées) et reproductible (refaire le test n'en change pas le résultat) et qu'elle est de plus liée par corrélation à des effets comme le succès scolaire, qu'elle permet ainsi de prédire. Mais existe-t-il une fonction cognitive particulière, un mécanisme spécifique à l'œuvre dans notre cerveau, qui permettrait d'expliquer g et les corrélations entre les résultats aux tests de QI ? La réponse est plus compliquée et nous permet d'éliminer un premier suspect dans la question des capacités générales. En réalité, aucune étude n'a réussi à mettre en évidence l'existence d'un lien spécifique entre une faculté particulière (parmi les possibilités : attention, vitesse de traitement de l'information, mémoire de travail) et g ou QI. En revanche, il est évident que chaque test mobilise simultanément une multitude de fonctions cognitives.

Ceci est vrai aussi en dehors des tests : les fonctions cognitives sont fortement spécialisées, mais elles travaillent de façon entrelacée pour contribuer aux différentes facettes de nos comportements. La réponse à l'existence apparente d'une base unique pour l'intelligence est donc cet entrelacement et interdépendance des différentes fonctions cognitives.

« *Ce que l'on nomme l'intelligence générale est simplement une propriété émergente, résultant de la cascade de facteurs environnementaux, génétiques, cérébraux et cognitifs qui influent sur la performance aux différents tests.* » (Ramus, 2012, p. 6).

Donc, bien que différents talents existent, la corrélation statistique entre nos différentes capacités reste réelle, parce que celles-ci, souvent, opèrent et se développent ensemble en s'influençant réciproquement pendant notre développement, quitte à puiser dans différents aspects de la maturation cérébrale qui subissent des influences communes par les gènes et

l'environnement. Cependant, le problème de l'intelligence ne se pose pas vraiment lorsqu'il est question de capacité générale *versus* capacité locale à penser. La question concerne plutôt les heuristiques que nous mobilisons pour résoudre différents types de problèmes, réfléchir à différents types de contenus. Certes, g pourrait (éventuellement) influencer la qualité de ces heuristiques, mais n'a rien à nous dire sur leur applicabilité plus ou moins large. Étudions donc les heuristiques.

4.2.2 Heuristiques générales : oui ou non ?

Une heuristique est une règle, une stratégie, simple, permettant de s'attaquer à un problème. Les années 1950-1970 ont été caractérisées par une certaine confiance en l'existence d'heuristiques générales (effectivement sous-tendues par le facteur g) permettant de s'attaquer à des classes larges de problèmes et à d'autres défis intellectuels. Un exemple d'heuristique de résolution de problèmes consiste à diviser le problème en sous-problèmes puis à résoudre un problème à la fois. Un autre exemple consiste à représenter un problème sous différents aspects. Et ainsi de suite. Durant l'âge d'or des heuristiques, ce type d'approches a donné lieu à des programmes d'intelligence artificielle portant le nom illustratif de « *General Problem Solver* » (GPS), créé par deux des pères des sciences cognitives, Allen Newell et Herbert Simon, avec J.P. Shaw. L'heuristique au cœur du GPS (à ne pas confondre avec le *Global Positioning System* qui fait preuve d'un tout autre type d'intelligence) est celle de l'analyse *means-end*. Un problème est défini par son état initial et final ; chaque étape de résolution du problème met en place un moyen qui permet de rendre l'état initial plus semblable à celui final. Après cette opération, les deux sont comparés et une nouvelle procédure recommence pour réduire ultérieurement la différence. Et ainsi de suite. Des principes généraux sont donc systématiquement appliqués, sans qu'il soit question de s'interroger sur la nature spécifique du problème à traiter. Au-delà de l'IA s'affirme la vision selon laquelle « bien penser » signifie posséder une série de bonnes heuristiques générales : pour la résolution de problèmes, pour bien mémoriser, pour inventer, pour prendre des décisions.

Le corollaire de cette vision est que chaque savoir local et base de données de connaissances jouent un rôle secondaire dans la pensée. Certes, personne ne nie que, pour devenir maître d'échecs, il est nécessaire de maîtriser les règles du jeu ou qu'un bon médecin a besoin d'une base de connaissances concernant la physiologie et la pathologie. Mais posséder et même maîtriser ces règles locales et bases de données est insuffisant pour expliquer les performances supérieure des maîtres d'échecs et des bons médecins, qui sont plutôt à ramener à leur maîtrise d'heuristiques générales.

Dans le cadre de cette vision, le chef d'État n'a en effet qu'à fournir à son maître d'échecs la base de connaissances appropriée pour que celui-ci, fort de ses heuristiques générales, se transforme en machine de guerre. Le problème de cette vision est qu'elle n'est pas supportée par les faits. Quatre domaines de recherche empirique minent plus particulièrement ses bases.

4.2.3 Atouts des experts

Et si nous analysions ce que les maîtres d'échec savent réellement faire, au-delà de leur capacité à gagner des parties ? Pendant les mêmes années que s'affirmait l'idée de « penseur général », des chercheurs ont étudié des aspects spécifiques des performances des experts, dont les maîtres d'échecs. Plusieurs questions ont alors été soulevées : les maîtres d'échecs font-ils preuve d'une mémoire époustouflante ? Sont-ils capables de garder le souvenir de matchs entiers et lointains dans le temps ? Cette attitude se traduit-elle en une plus globale capacité à se souvenir ? Est-il possible d'entraîner notre mémoire dans un domaine (les échecs ou autre) et d'en voir les bénéfices ailleurs ? Les résultats des études conduites à l'époque donnent des résultats plutôt décevants.

Par exemple, les maîtres d'échecs sont capables d'apprendre par cœur le chiffre faramineux de 50 000 configurations d'échecs différentes. Cependant, leur capacité à apprendre par cœur et à reproduire la position de pièces d'échecs en dehors des règles de l'échiquier n'est pas supérieure à celle de n'importe qui. En réalité, la capacité du maître d'échecs est supérieure seulement si les configurations à apprendre ont du sens et se rapportent aux connaissances et règles des échecs (de Groot 1965 ; Chase & Simon 1973) ; s'il dispose de plus de temps, il réussit légèrement mieux que les novices (Gobet & Simon 1996). La vision qui émerge alors de l'expertise va dans une direction opposée à celle des heuristiques générales. « *People who have developed expertise in particular areas are, by definition, able to think effectively about problems in those areas. Understanding expertise is important because it provides insights into the nature of thinking and problem solving. Research shows that it is not simply general abilities, such as memory or intelligence, nor the use of general strategies that differentiate experts from novices. Instead, experts have acquired extensive knowledge that affects what they notice and how they organize, represent, and interpret information in the environment. This, in turn, affects their ability to remember, reason, and solve problems.* » (Bransford et al. 2000)

Les experts utilisent une base de connaissances large et constituée de connaissances et savoir-faire relativement spécifiques. Par exemple, un scientifique de profession, un physicien, possède une base de connaissances concernant les lois et principes de la physique (Chi, Feltovich, Glaser 1981). Ces connaissances ne sont pas juste disposées quelque part ou indépendantes les unes des autres : elles sont faciles à mobiliser et sont fortement entrelacées entre elles, constituant un réseau de connaissances utilisables (Chi, et al. 1982).

Les experts sont capables de reconnaître rapidement les situations dans lesquelles ces connaissances s'appliquent, ils en perçoivent en quelque sorte la structure « profonde ». Par exemple, confronté à un problème concernant un plan incliné, un physicien y verra immédiatement à l'œuvre des lois du mouvement. Le maître d'échecs, de son côté, perçoit les configurations sur l'échiquier avec leur signification, pas simplement comme des pièces disjointes : il n'a donc pas à chercher à les relier par un processus abstrait, mais en quelque sorte les voit directement comme liées par un pattern (Bransford et al 2000[220]). Les experts raisonnent vers l'avant : ils utilisent leurs connaissances et principes pour y intégrer les constats et les nouveaux problèmes rencontrés (Larkin et al. 1980). Ils semblent avoir recours à la métacognition : ils évaluent leur niveau de compréhension et se demandent si les connaissances à leur disposition suffisent pour résoudre le problème actuel. Ils peuvent se poser la question de ce qui compte comme expertise dans leur domaine (Bransford, Brown & Cocking 2000). Bransford, Brown & Cocking (2000) synthétisent les spécificités des experts en six points[221]. De plus, ils considèrent que la notion d'expert est importante pour l'apprentissage, car elle permet de montrer ce qu'un apprenant peut atteindre en termes de capacités à penser dans un domaine.

À la figure de l'expert s'oppose celle du novice, qui se comporte de façon opposée : il manque d'une base de connaissances large et profonde, manque des habitudes qui permettent de reconnaître rapidement une structure profonde, reste « collé » à la surface du problème et tend à raisonner d'abord sur le nouveau problème et ses inconnus pour aller seulement après chercher dans ses connaissances celles qui pourraient l'aider à le résoudre. Ces deux figures émergent d'une grande variété de domaines disciplinaires : mathématiques, programmation, médecine, science, et naturellement échecs. Naît ainsi une théorie de l'expertise : les experts « voient » les choses différemment ; face à un même problème, ils ont accès rapidement à des connaissances entrelacées dans lesquelles une nouvelle pièce d'information trouve son sens ou peut être interrogée à la lumière de principes généraux (Chi, Glaser, Farr 1988).

La notion d'expertise reçoit aussi une nouvelle dimension dans le domaine de l'intelligence artificielle où, à partir des années 1970, les « systèmes experts » commencent à supplanter les « *general problem solvers* ». En médecine, par exemple, les chercheurs s'aperçoivent des effets puissants produits par des bases de connaissances larges. Ils commencent à considérer

les heuristiques générales comme des « méthodes faibles », qui coûtent peu en termes de base de données mais qui comportent aussi des résultats faibles. Les nouveaux systèmes experts se donnent pour objectif de simuler les réelles capacités des experts, à partir des caractéristiques décrites ci-dessus et donc de bases de données massives par exemple sur différents diagnostics médicaux.

4.2.4 Le cerveau n'est pas un muscle

D'autres recherches contribuent à défaire le mythe de l'entraînement global de capacités générales. Elles démontent en particulier le mythe selon lequel le cerveau serait une sorte de muscle que nous pourrions entraîner pour en améliorer les capacités de façon générale, indépendamment du contexte et des contenus. C'est le cas de la mémoire à court terme.

Des chercheurs ont demandé à un sujet de garder en mémoire autant de chiffres que possible, pour ensuite les répéter. Cette tâche rappelle celle de mémoriser les chiffres d'un numéro de téléphone avant de le noter quelque part. Au quotidien, nous nous rendons tous compte que cette capacité est limitée (aux alentours de sept éléments). Ce ressenti est confirmé par des recherches conduites à l'aube des sciences cognitives (Miller 1956). Il est également partagé par le sujet de l'expérience que nous allons relater. Sauf que, au bout de deux ans (au rythme de deux à cinq jours d'entraînement par semaine, deux heures par jour), le sujet en question – connu par ses initiales SF – est capable de réciter 81 chiffres d'affilée (Chase & Ericsson 1982, Ericsson & Chase, 1982). Pour faire bref avec une histoire longue, son long et coûteux entraînement n'a en réalité pas développé sa mémoire à court terme en tant que faculté cognitive globale, mais seulement sa capacité à mémoriser des chiffres. Pour la même tâche de mémorisation mais avec des lettres, ses performances retombent à sept éléments mémorisés ! Interrogé sur son entraînement, SF a révélé avoir utilisé des stratégies afin de compacter en un seul « objet » mental (un *chunk*, terme technique utilisé) plusieurs chiffres : passionné par les Jeux olympiques, il a associé des ensemble de chiffres à un événement, une course par exemple, et en a ainsi réduit l'espace en mémoire à court terme.
Cette étude est souvent utilisée pour appuyer l'idée selon laquelle la capacité à intégrer de nouvelles informations dans notre espace mental dépend des connaissances que nous possédons et de notre capacité à les utiliser à bon escient pour penser. En réalité, pour penser, nous utilisons notre mémoire de travail à court terme– l'espace dans lequel nous manipulons mentalement informations et représentations – mais également notre mémoire à long terme.

4.2.5 Savoir et savoir appliquer

Il est trivial de souligner que nous pensons toujours à quelque chose, à un certain contenu, et que ce contenu appartient à un domaine précis. Ainsi, les connaissances dans un domaine influencent-elles nos prestations, la compréhension par exemple (Recht & Leslie 1988 ; Willingham 2007, 2019) ? De très nombreuses études montrent que les connaissances antérieures (le niveau d'expertise) sont des prédicteurs qui expliquent les différences de performances entre sujets, et ceci indépendamment du facteur développement (âge des sujets[222]). Il est en revanche moins trivial de montrer que des processus de raisonnement différents sont mis en jeu selon les domaines. Comme exemple d'influence des contenus sur les capacités de raisonnement, citons la tâche de sélection de Wason : « *Quatre cartes comportant un chiffre sur une face et une lettre sur l'autre sont disposées à plat sur une table. Une seule face de chaque carte est visible. Les faces visibles sont les suivantes : D, 7, 5, K. Quelle(s) carte(s) devez-vous retourner pour déterminer la véracité de la règle suivante : si une carte a un D sur une face, alors elle porte un 5 sur l'autre face ? Il ne faut pas retourner de carte inutilement, ni oublier d'en retourner une.* » (Wikipedia : la tâche de sélection de Wason). Moins de 25 % des personnes soumises à ce test donnent la bonne réponse (D & 7). De plus, l'entraînement formel des capacités logiques ne semble pas avoir d'impact sur les prestations au test. Cependant, celles-ci tendent à s'améliorer significativement si la même

tâche est présentée sous un « aspect moral », donc le contenu change mais pas la logique appliquée. « *Quatre personnes sont en train de boire dans un bar et vous disposez des informations suivantes : la première boit une boisson alcoolisée, la seconde a moins de 18 ans, la troisième a plus de 18 ans et la dernière boit une boisson sans alcool. Quelle(s) personne(s) devez-vous interroger sur leur âge ou sur le contenu de leur verre pour vous assurer que tous respectent bien la règle suivante : si une personne boit de l'alcool, elle doit avoir plus de 18 ans.* » Lorsque l'identification de l'erreur logique révèle un tricheur, les prestations s'améliorent et le succès dans la résolution monte à 85 % (Cosmides & Tooby 1992) !

Prenons maintenant le cas d'une personne ayant appris qu'il est important, face à un problème, à un choix, d'opérer, d'avoir recours à des stratégies métacognitives particulières : cette personne aurait suivi un cours de pensée critique. En présence d'un nouveau problème ou d'une nouvelle circonstance, elle saura qu'il ne faut pas céder au biais de confirmation, ou à l'argument d'autorité, qu'il est important de suspendre son jugement pour évaluer équitablement différentes options, douter, se demander si l'argument est valide et s'il est en plus appuyé par des faits (donc possiblement vrai). Nous pouvons également savoir que, face à un problème complexe, une stratégie efficace consiste à le diviser en plusieurs sous-problèmes plus petits. Mais lesquels ? Comment opérer la division ? Jusqu'à quel niveau de division arriver ? Pour répondre à ces questions pratiques ou théoriques, il est utile de posséder une certaine expérience de problèmes analogues et des connaissances sur les constituants du problème… L'importance des contenus pour l'acquisition de capacités supérieures de pensée ne se limite donc pas au fait que la pensée a besoin d'un contenu de connaissances sur lequel s'exercer : cette capacité s'améliore en raison de la quantité de connaissances possédées et que mieux nous connaissons un certain domaine, mieux nous sommes capables de le penser de manière efficace et critique.

Le problème des heuristiques générales, déconnectées d'un domaine d'application et de connaissances de domaine, est donc celui de leur application. Même si nous suivons les règles adaptées pour résoudre un problème, si nous ne savons pas quand comment et de quelle manière spécifique les mettre en pratique, cela ne fonctionne pas. Les études sur l'expertise et celles sur la mémoire à court terme revalorisent l'importance, dans le raisonnement, des connaissances (la base de données) et, dans la capacité à penser, de la mémoire, en tant que fonction qui permet de stocker ces connaissances. Penser de façon efficace, comprendre en profondeur, trouver une solution experte sont toutes des capacités qui dépendent fortement de capacités et de contenus de connaissances qui sont spécifiques à un domaine.

4.2.6 Stratégies (apprendre à apprendre)

Une autre partie de la littérature en psychologie cognitive et de l'éducation utile pour notre discussion fait référence aux concepts de stratégie et d'apprendre à apprendre (Fayol & Monteil 1994[223]). La notion de stratégie est particulièrement développée dans la littérature sur la mémoire, et les stratégies de mémorisation : relire, répéter, préparer des résumés en sont des exemples (Bjorklund & Harnishfeger 1990). Il existe désormais une vaste littérature d'études scientifiques concernant les techniques les plus efficaces pour mémoriser des connaissances – la pratique du test, celle de l'apprentissage espacé, de la recherche de *feedback* et, plus généralement, de l'engagement actif et avec effort dans l'apprentissage (Brown, Roediger & McDaniel 2014). Une stratégie présente des caractéristiques précises (qui l'apparentent à des heuristiques telles que nous les avons décrites auparavant, du moins les stratégies d'ordre général[224]). Les stratégies se distinguent notamment des procédures élémentaires et des algorithmes car :

- le sujet dispose de multiples procédures et peut choisir entre elles ;

- il sélectionne les procédures sur la base de critères comme le but, les circonstances, la connaissance de ses possibilités.

Ainsi, une stratégie n'est pas une réponse « automatique » à une situation donnée, imposée par le fait qu'elle est la seule disponible ou la seule évoquée par la situation. De plus, elle n'est pas mise en pratique de manière automatique, selon un schéma préconstitué d'étapes. Le sujet fait un choix et ensuite guide son action et l'évalue d'une manière qui exige de l'attention. Mettre en place une stratégie est donc coûteux cognitivement, mais le gain est celui de l'adaptabilité et de la flexibilité par rapport aux circonstances et aux sujets. La capacité de mettre en place une stratégie fait donc partie des capacités de pensée de haut niveau, comme la pensée critique (Resnick 1987). Selon un constat assez général, les stratégies, une fois acquises, ne sont pas systématiquement réutilisées. Par exemple, l'élaboration de résumés pour mieux apprendre est une stratégie qui peut être enseignée ; les élèves l'utilisent au cours de la leçon mais l'abandonnent dès la séance suivante (Fayol & Monteil 1994).

Une série de difficultés s'interposent entre le fait d'avoir appris une stratégie générale et celui de l'utiliser dans la vie – même scolaire : la difficulté de percevoir la valeur ajoutée de la stratégie, en l'absence d'un *feedback* approprié et de l'information sur la pertinence de son utilisation (déjà qu'il faut concevoir que ce que nous faisons peut avoir un impact sur notre réussite, que celle-ci n'est pas une sorte de destinée mais dépend de l'effort) ; le coût cognitif de l'effort ; enfin, le rôle des connaissances préalables[225]. Au point que l'apprentissage de procédures serait facilité lorsque leur introduction et l'entraînement à les utiliser se feraient sur des contenus maîtrisés, au niveau des connaissances, par le sujet. Les savoirs préalables peuvent ainsi constituer un obstacle mais aussi un atout pour l'acquisition de stratégies car ils évitent la surcharge cognitive.

4.3. Apprendre à penser de façon (plus) critique : un faux espoir ?

Il n'existe pas de consensus autour du fait que les interventions visant l'EC sont efficaces. Cette absence est motivée par deux raisons principales :

- les différents auteurs n'utilisent pas les mêmes critères pour définir ce qui constitue un résultat positif ;
- les auteurs n'utilisent pas la même définition de l'EC.

À ces deux difficultés s'ajoute le fait que les interventions évaluées sont très différentes entre elles : en durée, en typologie d'activités, mais aussi en objectif et en vision théorique de ce qu'est l'EC. Toutes ces difficultés rendent les méta-analyses quantitatives et les revues systématiques qualitatives de la littérature difficiles à mener, et en réduisent la pertinence.

4.3.1 Le pessimisme de Daniel Willingham

Le psychologue cognitif Daniel Willingham a effectué une rapide analyse de l'efficacité des programmes les plus connus d'enseignement de l'esprit critique *per se* (Willingham 2007). Il synthétise les caractéristiques communes aux méthodes plus diffusées en trois points :

- elles présument l'existence de capacités (*skills*) pouvant être pratiquées indépendamment du contexte et du contenu, donc l'enseignement a lieu en dehors d'un cadre disciplinaire ;
- certaines ont une durée longue (trois ans, plusieurs heures d'instruction par semaine) ;

- elles utilisent toutes des exemples de pensée critique et ensuite demandent d'appliquer ; certaines utilisent des problèmes abstraits (*Ruven Feuerstein Instrumental Enrichment*), d'autres des histoires à mystères (*Martin Covington Productive Thinking*), la discussion de groupe de problèmes quotidiens (*Edward De Bono Cognitive Research Trust* : CORT).

Les études qui mesurent les effets de ces interventions ont plusieurs limites :

- les étudiants sont évalués seulement une fois, donc nous ne pouvons pas savoir si les effets perdurent ;
- il n'y a pas de groupe témoin, ou le groupe témoin ne mène pas une activité alternative (groupe passif) ;
- il n'y a pas de mesure de transfert à des situations réelles ou à des situations différentes de celles utilisées dans le cadre de l'instruction ;
- seule une petite partie de ces études ont été soumises au processus scientifique de publication et de validation par les pairs ;
- quand nous évoquons les effet positifs de ces méthodes, nous ne parlons pas nécessairement de la même chose ou de quelque chose qui a spécifiquement trait avec l'esprit critique[226].

En dépit de toutes ces difficultés et limites, les analyses de la littérature ou les textes sur l'esprit critique indiquent souvent que les interventions éducatives pour l'esprit critique ont des effets positifs[227].

En s'inspirant de la littérature que nous avons citée ci-dessus, Willingham (2007) rejette en outre l'idée selon laquelle la pensée peut être considérée comme une capacité (*skill*) analogue à celle qui permet de faire du vélo. La pensée est fortement dépendante des contenus de la connaissance. Rappeler à quelqu'un d'utiliser son esprit critique, ou de penser à chercher les hypothèses alternatives, etc. ressemble à la situation où nous disons à quelqu'un : « Il faut faire attention[228]. » Ce type d'exhortation peut servir de « rappel » même quand nous ne savons pas à quoi faire attention et comment cela risque de ne pas amener au comportement souhaité. La dépendance de la pensée au contenu se révèle aussi dans la difficulté à se détacher des aspects superficiels, concrets d'un problème (son contenu factuel) pour chercher à percevoir sa « structure profonde ». Quand nous cherchons à comprendre un nouveau problème, nous utilisons les connaissances stockées dans notre mémoire ainsi que le contexte. Ceci rend la compréhension plus rapide, mais l'ancre au contenu contextuel. Ces considérations amènent Daniel Willingham à affirmer qu'enseigner – de façon générale – à penser ou à penser de façon critique est impossible[229].

N'existe-t-il aucun espoir de transférer des acquis d'un contexte à un autre – y compris des acquis « généraux » comme la capacité à résoudre un certain type de problème ? Willingham en indique deux : la familiarité avec le contenu profond et le fait de savoir qu'il faut regarder ce contenu profond[230]. La familiarité dépend d'une pratique répétée, et automatisée[231]. La deuxième stratégie est métacognitive : il s'agit de penser à regarder en mémoire s'il existe quelque chose de semblable[232]. Lorsque cela ne vient pas automatiquement grâce à la familiarité, il faut au moins avoir en tête le principe de le faire. Pour vraiment l'implémenter, des connaissances et de la pratique sont nécessaire. Sans cela, nous savons ce que nous devrions faire mais nous ne savons pas comment le faire. Nous pouvons en conclure que la capacité de penser de manière plus critique dépend fortement du contenu.

Cela est également vrai pour des formes plus spécialisées de pensée « experte », comme la pensée scientifique. Même dans le cas de la pensée scientifique, le succès dépend non seulement du fait de connaître les procédures, les stratégies, mais aussi de savoir quand les appliquer et comment[233]. La conclusion finale est que, bien que nous détenions des capacités

naturelles, comme celle de raisonner sur les causes et les probabilités conditionnelles (une capacité qui est, sous une forme intuitive, limitée dès la petite enfance), nous pouvons continuer à commettre des erreurs dans l'usage des probabilités conditionnelles et dans le raisonnement causal même en présence de capacités plus sophistiquées et d'un certain nombre de connaissances[234]. L'enseignement de l'esprit critique repose donc, en partie, sur le fait d'apprendre aux élèves des stratégies pour « mieux penser », et en grande partie sur comment et quand déployer ces stratégies[235].

4.3.2 L'optimisme de Diane Halpern

Diane Halpern, également psychologue cognitive, est beaucoup plus optimiste quant à la possibilité d'éduquer l'esprit critique et au succès des méthodes existantes (Halpern 2013[236]). Elle cite notamment des études conduites par d'autres chercheurs ou groupes, comme une méta-analyse du *Thinking Skills review Group* (Higgins et al. 2005). Elle inclue des interventions qui impliquent de la métacognition sur ses processus d'apprentissage et de pensée, des activités liées à la créativité mais aussi à des capacités de raisonnement (verbal, spatial, logique).

Halpern cite aussi des études ayant évalué l'impact d'interventions spécifiques concernant les capacités de pensée, à grande ou à petite échelle. Une évaluation nationale au Venezuela, menée par le gouvernement avec l'assistance de Harvard sur des élèves de treize ans en moyenne, mesure ainsi les effets d'une intervention à grande échelle, randomisée et contrôlée (Herrnstein et al. 1986). Van Gelder (2001) évalue son propre programme centré sur l'apprentissage de stratégies d'argumentations, programme dédié aux étudiants universitaires. Marin & Halpern (2010) analysent l'impact différentiel d'interventions explicites et implicites de capacités de l'EC pouvant être mesurées par le *Halpern Critical Thinking Assessment* (HCTA) : la cible de l'intervention et de l'évaluation sont des élèves adolescents en études secondaires supérieures. Les résultats sont comparatifs : l'instruction explicite fonctionne mieux que celle implicite (Heldsdingen et al. 2010). Il existe également un test réalité aux Pays-Bas relatif à une intervention, comparée avec aucune intervention. Sont aussi cités : Facione (1991) ; Facione (2000) ; Lehman, Lempert, Nisbett (1998) qui testent leurs interventions *via* des tests à choix multiples et des mini-scénarios. Pour Halpern, le groupe d'études les plus significatives, notamment pour les questions de transfert, sont celles menées par Nisbett et reportées dans Nisbett, Jepson, Krantz (1993), Nisbett et al. (1993), Nisbett et al. (1987) et Fong, Krantz, Nisbett (1986). (Voir aussi: Nisbett 2013)
Les interventions de Nisbett sont notamment centrées sur des aspects comme la logique, le raisonnement causal et probabiliste, l'analyse coût-bénéfices. Nisbett utilise des situations réalistes à la fois dans le cadre de sa pédagogie de l'esprit critique et dans le cadre de ses tests. Le plus « extrême » point de vue sur la capacité de mesurer le transfert à des situations de la vie réelle est celui conduit par téléphone sans que les sujets de l'expérience ne sachent qu'il sont en train de passer un test d'esprit critique. À partir de ses résultats, Nisbett considère que l'esprit critique est une capacité (*skill*) qui s'apprend et se transfère à de nouveaux contextes et contenus. Kosonen et Winne (1995) montrent des résultats positifs – du collège à l'université – pour des apprentissages de règles de raisonnement appliquées dans des contextes différents. La conclusion qu'en tire Halpern est que le transfert est possible modulo ce type de stratégie : répétition dans plusieurs contextes et enseignement des règles. Elle considère que le succès de l'enseignement de la pensée critique devrait indiquer la voie à l'enseignement plus général lorsque nous voulons obtenir un transfert. Au contraire, l'instruction non spécifiquement dédiée à améliorer l'esprit critique ne semble pas avoir d'effet sur celui-ci[237].

Cependant, Halpern pourrait être trop optimiste si nous considérons les difficultés du transfert, à moins d'admettre le rôle du contenu dans l'exercice de la pensée et des déterminantes innées dans notre manière de penser qui ne se laissent pas facilement effacer : ce qu'on appelle nos « biais ». Halpern propose également une vision très « haut niveau » de

ce que l'esprit critique devrait permettre de faire. Sa vision implique non seulement l'amélioration de certaines de nos capacités (analyse du langage, utilisation de probabilités, plausibilité, compréhension de ce qui compte comme preuve...) mais aussi la mise en place de capacités de « résistance » à nos biais et tendances naturelles. Et ceci dans des conditions et contextes « naturels » qui impliquent un temps limité, du stress, des émotions, des motivations sociales[238]…

Le procédé de Halpern pour obtenir une amélioration des capacités qui composent l'esprit critique se base sur quatre stratégies éducatives :

- enseigner l'esprit critique de façon assumée, sans espérer qu'il se développera seul dans le cadre d'une instruction « normale » ;
- enseigner de façon explicite les stratégies ;
- faire pratiquer dans une variété de contextes et sur une variété de contenus les mêmes stratégies[239] ;
- utiliser des outils pédagogiques qui se sont montrés utiles pour d'autres apprentissages, notamment la stratégie du "test" : se tester pour mieux mémoriser, à la fin comme au milieu de l'apprentissage. Cette stratégie peut être pratiquée pour multiplier la pratique dans des divers contextes et avec des contenus différents.

En dépit de la différence entre le pessimisme de l'un et l'optimisme de l'autre, ces quatre indications sont communes à Willingham et à Halpern, qui de toute manière baignent dans un même champ de connaissances concernant l'apprentissage et le transfert.

Les études citées par Halpern (2013) apportent en réalité des preuves faibles de l'éducabilité de l'EC, grâce aux méthodes existantes. Voyons quelques-unes se ces études plus en détail. Nous nous arrêterons sur la méta-analyse de Higgins et al. (2005), sur l'étude expérimentale la plus large conduite (Herrnstein et al. 1986), sur l'étude conduite par Halpern et elle-même (Marin & Halpern 2010) et sur les études de Richard Nisbett, qui semblent être les plus prometteuses en termes de transfert. Ensuite nous passerons à analyser une méta-analyse plus récente (Abrami et al. 2015).

4.3.2.1 Une méta-analyse

La méta-analyse, quantitative, analyse publiée en 2005 par Higgins et al., fait partie des études systématiques d'impact du EPPI Centre de UCL. Le EPPI Centre joue un rôle de promotion de l'*evidence-based* en éducation et, plus généralement, dans le domaine des politiques sociales. Il sert aussi d'organisme de production d'études et de méta-analyses pour le EIPPEE européen. Nous profitons de cette étude pour souligner que les méta-analyses sont toujours à privilégier dans le cadre d'une appréciation de l'impact et que la motivation pour l'analyse est que l'EC constitue un objectif pédagogique dans les programmes en Angleterre et au Pays de Galles. Higgins et al. (2005) ne se limitent pas à prendre en considération l'EC, mais les capacités de pensée, de planification et la métacognition au sens général[240]. Ainsi leur revue de la littérature inclut des programmes tels que « *instrumental enrichment* » (Feuerstein et al., 1980), philosophie pour enfants (Lipman et al., 1980), « *cognitive acceleration through science education* » (CASE) (Adey et al., 1989), « *Somerset thinking skills* » (Blagg 1988) et des programmes utilisant des méthodes centrées sur le dialogue, l'argumentation, etc.

Dans la méta-analyse sont comprises 29 études. Elles ont été sélectionnées en suivant les critères du EPPI Centre. Cependant, ceux-ci sont assez larges (en effet, seuls 13 sont randomisés[241]). Les résultats de la méta-analyse, considérés comme positifs par les auteurs, se réfèrent à des mesures d'impact sur des capacités cognitives diverses, certaines n'appartenant pas au domaine de l'esprit critique (par exemple, les matrices de Raven : taille de l'effet 0,62, typiquement inclus dans les tests de QI mais sans relation spécifique avec la pensée critique

ou d'autres formes d'analyse des informations). Les auteurs indiquent aussi une amélioration des résultats scolaires, qui peut être mise en relation avec des programmes spécifiquement adressés aux compétences en sciences, comme CASE, mais difficilement mises en relation avec les interventions pour améliorer la pensée (les résultats scolaires dépendent largement des connaissances). Il existe une grande hétérogénéité entre interventions (certaines concernent la manière de favoriser des méthodes de penser optimistes). Au fond, ces résultats concernent plutôt des méthodes d'enseignement que l'enseignement de l'esprit critique – et leurs résultats sont en effet compatibles avec ceux obtenus par Hattie et Marzano à propos des méthodes impliquant de la métacognition (Hattie, Biggs & Purdie 1996 ; Hattie & Donoghue 2016 ; Marzano 1998 ; Higgins et al. 2016). L'hypothèse qu'elles pourraient à la limite appuyer est donc celle selon laquelle il est possible d'obtenir des résultats positifs quand nous nous adressons à l'éducation de capacités de pensée et non à des contenus disciplinaires, mais pas dans la direction de l'EC. Nous trouvons également beaucoup d'hétérogénéité dans l'âge des sujets des différentes études et dans les types de mesures (tests standardisés ou non). Il est donc préférable d'être très prudents avec l'interprétation de ces résultats. La méta-analyse met aussi en évidence l'existence de biais de publication, qui suggère celle d'études publiées avec résultats négatifs.

4.3.2.2 Une large intervention

Herrnstein et al. (1986) décrivent les résultats d'une intervention sur environ 400 étudiants de collège (13 ans) au Venezuela. Cette intervention a la particularité de ne pas porter sur des contenus curriculaires mais sur des capacités comme le raisonnement « scientifique » (observation et classification), l'utilisation critique du langage, l'argumentation, la créativité, la résolution de problèmes, la prise de décision. L'intervention se compose d'un total de 100 leçons dont environ 60 utilisées en classe. Les élèves sont tous d'origine LSES. L'intervention dure un an, et est systématique : trois leçons se déroulent sur les quatre jours d'école, le quatrième étant utilisé pour la révision. Les cours sont donnés par des enseignants externes sélectionnés et préparés pour la tâche. Une batterie de tests est administrée aux élèves des classes test et des classes contrôle (contrôle passif) en prétest et en post-test. La batterie comprend des tests standardisés et non pour un total d'environ quatre heures et demi de passation. Les tests non standardisés mesurent les capacités travaillées en changeant de contenu. Les élèves du groupe expérimental sont meilleurs dans tous les pré-tests que les élèves du groupe contrôle. L'effet global est considéré comme positif car les élèves du groupe expérimental progressent plus, notamment dans les tests non standardisés, que les élèves du groupe témoin. Herrnstein et al. (1986) posent correctement la question de ce qui peut être considéré comme un succès pour un tel type d'initiative et de comment distinguer entre effets produits par les contenus de l'interventions et effets dus à l'effet « intervention » ou effet Hawthorne. En effet, ils remarquent un effet établissement et enseignant.

4.3.2.3 Une étude d'impact d'une méthode spécifique

Marin & Halpern (2010) ont évalué l'impact, sur des étudiants d'études secondaires, d'une intervention de type explicite et pratique conforme aux quatre piliers de l'enseignement de l'esprit critique selon Halpern : travail sur les attitudes et encouragement des étudiants à la réflexivité ; enseignement et pratique de capacités liées à l'esprit critique ; entraînement pour le transfert, en identifiant des situations concrètes où adopter les stratégies apprises ; métacognition comme réflexion sur ses processus de pensée. L'intervention est comparée à une intervention implicite (enseignement d'un cours de psychologie cognitive avec des contenus communs mais implicites) et avec un groupe passif en « attente » de son cours. Respectivement, l'étude a porté sur 28 + 18 + 24 étudiants. L'intervention a consisté en un pré-test, quatre leçons, puis un post-test. Elle s'est déroulée en trois semaines, au rythme de deux leçons de deux heures et demi par semaine. La participation était volontaire et a eu lieu après la classe ou le samedi. Étaient présents pour chaque leçon un instructeur, deux aides

chercheurs, un technicien. L'intervention se basait sur un support informatique que l'élève utilisait seul et du matériel pour l'utilisation en classe avec l'instructeur. Les deux groupes, explicite et implicite, ont progressé entre pré-test et post-test, mais le groupe expérimental est celui qui a le plus progressé.

Pour pallier les limites de l'étude (petit nombre d'effectifs, biais d'auto-sélection…), celle-ci a été refaite avec six leçons directement données dans le cadre des heures scolaires, deux fois par semaine pendant six semaines (40 + 38 + 30 étudiants pour les groupes expérimental, témoin actif, témoin passif). Le groupe passif n'a pas réalisé de post-test après quelque temps. Les résultats montrent des gains de performance pour les étudiants du groupe explicite mais pas pour ceux du groupe implicite.

4.3.2.4 Des études avec mesure de transfert à distance

Le psychologue Richard Nisbett occupe une position diamétralement opposée à celle de Willingham concernant la capacité d'apprendre les capacités générales abstraites d'un contenu particulier. Il soutient que sa vision a été modifiée après avoir entrepris des études avec Fong et Kranz pour démontrer justement cette limite. En étudiant l'impact de l'instruction aux statistiques, il se serait rendu compte que, en réalité, ce type d'instruction a des effets généraux qui vont au-delà des contenus enseignés. D'autant plus que les effets seraient déjà visibles avec des entraînements relativement courts (Nisbett 1993). Nisbett a testé ses effets dans des situations écologiques : en appelant les sujets à la maison et en prétendant mener une enquête, il a posé des questions relatives aux compétences statistiques et à leur application dans des cas différents de ceux ayant fait objet d'une instruction. Cependant, les questions posées pouvaient rappeler les notions de statistiques travaillées, car elles en comportaient le langage mathématique. En outre, les participants du cours de statistiques répondaient aux questions de l'enquête peu de temps après avoir terminé ce cours. Nisbett a reproduit ses études sur d'autres types de contenu, comme les règles pour évaluer la causalité, l'analyse coûts-bénéfices, etc. Selon lui, dans des domaines où nous possédons des règles intuitives (et imparfaites) de résolution, il est possible de faire apprendre des systèmes de règles plus sophistiqués (les règles des statistiques, de l'économie) et l'apprenant peut ensuite les appliquer de façon générale. Cependant, Nisbett reconnaît que des différences de performance existent dans le transfert d'une règle générale à des situations concrètes. Ces différences peuvent être attribuées à la transparence du problème. Certains problèmes laissent en effet voir à travers leur contenu superficiel le système de règles nécessaire à leur résolution ; d'autres ne le font pas et donc appliquer les règles devient difficile dans ces cas. Nisbett reconnaît également que les systèmes de règles qui permettent ce type de traitement général sont relativement peu nombreuses : les règles qui permettent d'identifier la bonne cause, les schémas contractuels dans la sphère sociale, les règles statistiques comme la loi des grands nombres[242].

4.3.3 Une nouvelle synthèse : Abrami et al. (2015)

Récemment, Abrami et al. (2015) ont entrepris de synthétiser les différentes études concernant l'impact des méthodes pour l'enseignement de l'esprit critique. Leur recherche a été effectuée en 2003, puis mise à jour plusieurs fois (la dernière en 2009). Au total, 684 études publiées ont été retenues pour l'analyse finale, car elles répondaient aux critères de pertinence établis pour la méta-analyse.

Une telle entreprise se heurte naturellement à plusieurs difficultés :

- la difficulté de définitions floues et non consensuelles du terme EC ;
- la difficulté représentée par la multiplicité des différentes méthodes, en particulier car celles-ci suivent des points de vue différents sur la nature générale ou plutôt spécifique

(au contenu) des capacités de penser et car elles ont des durées variables et sont plus ou moins intensives ;

- les évaluations de la réussite sont conduites à l'aide d'instruments divers, certains standardisés – comme le W*atson Glaser Critical Thinking Assessment*, le *Cornell Critical Thinking Test*, le *California Critical Thinking test* et le *Critical Thinking Disposition Inventory*, d'autres non. Parmi ces derniers, certaines formes d'évaluation sont créées par des enseignants, par les chercheurs ayant dessiné l'intervention ou par d'autres chercheurs, ou alors il s'agit d'échelles et tests non pensés pour évaluer l'esprit critique (y compris, les résultats scolaires) ;
- la difficulté représentée par la nature souvent quasi-expérimentale ou pré-expérimentale des études disponibles : ces méthodes sont testées dans des contextes « écologiques » (en classe), manquent de groupes de contrôle ou n'effectuent pas de distribution aléatoire des sujets.

En pratique, de nombreuses études publiées se limitent à comparer les résultats des pré-tests (avant intervention) à ceux des post-tests (après intervention) pour un même groupe de sujets, ce qui rend impossible de contrôler les effets test, les effets maître et les effets d'autres variables non identifiées sur le résultat final. Cependant, les auteurs de la méta-analyse ont décidé d'intégrer également dans leur sélection ces études quasi-expérimentales en considérant le résultat au pré-test comme équivalent à celui d'un groupe de contrôle. Il faut donc reconnaître que les résultats actuels, pris dans leur complexe, sont peu parlants et peuvent difficilement nous éclairer quant à l'efficacité d'enseigner l'esprit critique comme une capacité générale ou particulière.

Ces considérations faites, les résultats de la recherche sur l'éducation de l'esprit critique montrent-ils, dans leur complexité, des effets positifs ? La réponse est positive, avec plusieurs bémols directement issus des difficultés discutées ci-dessus. Il apparaît que les différentes méthodes examinées permettent de développer une ou plusieurs capacités identifiées comme faisant partie de la pensée critique, générales ou spécifiques à un contenu, selon les objectifs de l'étude et la méthode adoptée. Ceci à tous les niveaux éducatifs, à travers toutes les disciplines et pour toutes les stratégies pédagogiques adoptées.

Les stratégies pédagogiques sont classées en quatre grands groupes, selon une classification proposée par Ennis (1989) : enseignement général (les principes, capacités, dispositions sont enseignés de façon générale et pas en relation avec un contenu en particulier) ; enseignement immersif ; enseignement par infusion ; enseignement mixte. (Dans les enseignements immersif et par infusion, les contenus sont importants mais, dans le premier, les principes généraux de la pensée critique sont tout de même enseignés de façon explicite, comme des objectifs du cours.) D'autres modalités pédagogiques sont prises en considération de façon séparée, notamment : l'étude ou enseignement individuel ; l'enseignement par le dialogue, la discussion en groupe, le débat formel, et d'autres méthodes collaboratives ; la présentation de problèmes authentiques à résoudre, ancrés sur des contenus riches ; le coaching de la part d'un tuteur. Sont aussi comparées les interventions courtes, moyennes et étendues sur de longues durées. Toutes les stratégies, méthodes et durées donnent des résultats positifs par rapport aux groupes de contrôle sans instruction à la pensée critique (quand présents) ou aux pré-tests avant l'instruction à la pensée critique. Les impacts (taille de l'effet) sont cependant plutôt réduits. Des modalités d'instruction sont plus efficaces que d'autres pour le développement de capacités « générales » (non liées à un contenu disciplinaire spécifique) : d'un côté, les discussions guidées par l'enseignant et, de l'autre, l'exposition à des situations authentiques et à des contextes riches et situés, avec des exemples, en particulier lorsque sont utilisées deux méthodes : la résolution de problèmes et le jeu de rôle. Ces deux approches semblent particulièrement efficaces quand elles sont combinées et quand une forme de tutorat est ajoutée. Les études dans lesquelles les trois modalités sont combinées présentent en effet des tailles de l'effet significativement plus importantes que les trois prises individuellement

(le tutorat seul en revanche ne présente pas un effet positif plus important que les autres stratégies et méthodes mentionnées plus haut).

Naturellement, plusieurs questions restent ouvertes à la lumière de ces résultats. Le fait que toute intervention obtienne des effets ne fait que confirmer que « quelque chose vaut mieux que rien ». Mais nous ne pouvons pas savoir si l'effet se maintient ni s'il est transférable ailleurs que dans les tests effectués, donc s'il a un réel impact sur la vie et les attitudes de ceux qui ont reçu une forme ou une autre d'instruction à la pensée critique. En pratique, les études recensées permettent d'établir que les différentes formes d'instruction ayant été testées et dont les résultats ont été publiées permettent de produire des apprentissages qui sont au mieux des précurseurs d'une pensée critique génuine et transférable. Ceci est encourageant, mais invite encore plus à (se) poser la question du transfert dans toute forme d'instruction de la pensée critique. Il ne faut pas oublier que certaines méthodes semblent – à ce niveau – plus efficaces que d'autres, et qu'elles fournissent donc des indications sur les méthodes qu'il faudrait développer et tester prioritairement : présence d'argumentation, de contenus et d'exemples authentiques, accompagnement de la part de l'enseignant sous forme de tutorat.

4.4. Difficile ne veut pas dire impossible : réalité et difficultés du transfert

Si les capacités comme celle de penser ou de résoudre des problèmes sont générales, alors il devrait être facile de les transférer d'un contexte d'apprentissage et d'un contenu à un autre – voire, cela devrait arriver de façon automatique. Nous avons cité Edward Thorndike qui, au début du XIXe siècle, avait entrepris de tester dans les écoles si l'apprentissage du latin et des mathématiques avait un impact mesurable sur d'autres fonctions cognitives. Il avait répondu négativement. Encore une preuve que le cerveau n'est pas un muscle, et que les facultés mentales ne sont pas automatiquement développées par l'un ou l'autre apprentissage spécifique. Cette impression est confirmée par maintes études sur le transfert des apprentissages, qui peuvent se résumer ainsi : le transfert est la bête noire de l'apprentissage (Salomon & Perkins 1989). Pourtant, intuitions et expériences de la vie de tous les jours semblent suggérer que l'image de l'apprentissage comme complètement dépendant d'un contexte et d'un contenu (non seulement en phase d'acquisition, mais aussi dans celle de la réutilisation) est excessivement négative. Nous apprenons à faire du vélo et nous ne dépendons pas de notre premier vélo pour toute la vie. Nous apprenons à cuisiner des légumes et un effort moins important nous est requis pour apprendre à cuisiner du poisson. Nous apprenons à lire et, même nous n'avons pas la capacité de tout comprendre, en raison de nos connaissances, la lecture en tant qu'outil se généralise d'un contexte à un autre. Nous apprenons une manière de conduire des expériences ou à reconnaître une certaine typologie de raisonnements fallacieux et nous avons l'impression d'être capables de conduire des expériences dans d'autres domaines plus ou moins éloignés et de réutiliser ces typologies à partir de contenus assez éloignés entre eux.

Que fait un expert lorsqu'il est confronté à un problème atypique ? Il peut puiser dans sa base de données de connaissances et son savoir-faire. Il peut reconnaître la structure profonde du problème. Mais si la situation est réellement nouvelle, cela ne suffira pas. Plusieurs études montrent que, confrontés à un problème nouveau, les scientifiques ont recours au raisonnement par analogie. Ils cherchent ainsi à ramener la nouveauté sur un terrain connu en s'efforçant d'identifier dans la structure de la première des éléments qui la rapprochent de problèmes traitables par leurs moyens. Certaines études suggèrent donc que, dans ce type de défi, l'expert est forcé de retomber sur des heuristiques générales – comme celle d'avoir recours à un raisonnement par analogie –, mais sans oublier ses connaissances spécifiques. Et qu'il le fait afin de transférer ses compétences dans un domaine éloigné. (Le problème est, nous le mettrons de côté un instant sans l'oublier, qu'il est facile de tomber sur de mauvaises

analogies lorsque nous ne maîtrisons pas les contenus.) Une considération encore plus générale nous amène à douter de l'impossibilité du transfert. Nous nous trouvons, au fond, presque constamment confrontés à des situations nouvelles – ou du moins à des situations qui présentent à la fois beaucoup de similitudes avec celles rencontrées par le passé, mais aussi inévitablement des différences. Si nous ne pouvions en rien transférer nos apprentissages et savoir-faire, faire face à la nouveauté serait impossible. Or, non seulement nous répondons à la nouveauté, mais nous le faisons souvent, à tort, en ignorant sa part de nouveauté et en transférant des solutions inappropriées comme si nous nous trouvions face à une situation connue. En d'autres mots, nous appliquons souvent des stratégies qui nous sont disponibles, aveuglement et parfois inconsciemment, à une nouvelle situation. Alors pourquoi le transfert n'a-t-il parfois pas lieu ? Celle relative au transfert est probablement une question de degré, de qualité, et de méthode.

4.4.1 La notion de transfert est fondamentale pour l'apprentissage

La distinction entre transfert et apprentissage est floue, naturellement, car nous ne pouvons pas considérer un apprentissage comme tel s'il ne se manifeste pas par un niveau minimal de transfert d'une situation initiale à une nouvelle. La notion de transfert fait donc partie de façon essentielle de celle d'apprentissage, et nous ne pouvons pas nous donner comme objectif le deuxième sans viser du moins en partie le premier. La considération sur la centralité du transfert est d'autant plus importante lorsque nous nous posons la question des objectifs de l'éducation[243]. L'éducation permet de transmettre aux jeunes générations les savoirs et savoir-faire développés par leurs prédécesseurs, mais l'un des objectifs est aussi de leur permettre de faire usage de ces savoirs et savoir-faire, et de les améliorer, éventuellement. Pour cela, il est fondamental d'apprendre à transférer. Apprendre à transférer est aussi nécessaire pour remplir un autre objectif de l'éducation, celui de développer les capacités de pensée et de réflexion, pour améliorer la capacité des futurs citoyens à décider, faire des choix rationnels, interagir avec les autres et avec leur environnement de façon flexible. La question qui se pose est donc triple : pouvons-nous transférer des capacités, connaissances, attitudes relatives à l'esprit critique, qui auraient été apprises dans un certain domaine, dans un certain contexte, à un autre ? À la vie quotidienne ? De façon flexible ?

Quand Thorndike cherchait à tester la validité de l'hypothèse des disciplines formatrices (une thèse chère au psychologue Alfred Binet, inventeur du test de QI), il pensait en réalité à une autre théorie, la sienne (Thorndike & Woodworth 1901, a, b). Selon lui, le transfert entre une situation initiale, d'apprentissage, et une nouvelle situation d'application est possible mais seulement à condition qu'il existe des éléments communs aux deux situations (*common element theory*). Dans cette formulation générale, la théorie est largement acceptée. La distinction classique et répandue entre transfert proche et lointain remonte à cette idée : est dit proche le transfert entre situations qui partagent des éléments identiques, ou qui présentent beaucoup de similitudes, y compris des similitudes de surface. Par exemple, nous pouvons apprendre à dessiner au crayon puis transférer cette compétence dans des domaines progressivement plus éloignés : la craie, l'aquarelle, l'huile… Certaines compétences peuvent même être transférées sur la photo. Ce type de transfert a plus de chances de réussir que le transfert lointain – qui d'ailleurs n'est pas garanti (Sala & Gobet 2017).

Le problème avec cette théorie est représenté – comme nous allons le voir par la suite – par l'ambiguïté du terme « élément » et par la difficulté de définir de manière unique la distance entre éléments, c'est-à-dire le type de similitude entre situations permettant le transfert. L'attention s'est notamment concentrée sur les éléments « internes » aux situations d'apprentissage et situations cibles sur lesquelles effectuer le transfert : objets, actions, événements présents dans l'une et dans l'autre, relations logiques entre ces éléments internes (Gyck & Holyoak 1980, 1983 ; Gentner et al. 2003). Dans les années 1980, par exemple, Anderson reprend la théorie de l'identité des éléments et soutient que le transfert dépend de l'existence de similitudes entre les connaissances et opérations cognitives qui sont requises

par la tâche d'apprentissage et celle cible (Singley & Anderson 1989 ; Anderson, Reder, Simon 1996).

4.4.1.1 Transfert et analogie

Un médecin est confronté à un cas difficile : il doit soigner un patient qui souffre d'un cancer de l'estomac ; le patient mourra si la tumeur n'est pas détruite, mais l'opération est impossible. Il peut employer des radiations assez puissantes pour se débarrasser de la tumeur, mais leur passage à travers les tissus sains les détruira aussi. Si la puissance maximale n'est pas utilisée, la thérapie est inefficace. Comment détruire la tumeur à l'aide de la radiation sans détruire en même temps les tissus sains du patient ?
Ce problème est connu comme le « problème de la radiation de Dunker », du nom du psychologue qui l'a introduit en 1945, dans le cadre d'une réflexion sur la résolution de problèmes (Dunker 1945). Présentées de la sorte, les chances de résoudre le problème sont plutôt faibles. D'autres psychologues, l'ayant repris par la suite, les ont estimées dans leurs études à environ 10 % (Gick & Holyoak 1980, 1983). Ce chiffre peut nous sembler trop bas, parce que nous avons l'impression que le problème n'est pas si difficile. Notre intuition est que nous sommes devenus familiers avec les traitements contre le cancer et que certaines solutions nous viennent spontanément à l'esprit… parce que nous les connaissons. La connaissance des mécanismes de fonctionnement des radiations permet en effet de résoudre ce problème efficacement (Helfenstein & Saariluoma 2006). Un autre facteur qui augmente significativement les chances de succès concerne le fait de pouvoir disposer d'un cas analogue résolu, comme l'exemple suivant.

Une petite nation est dominée par un dictateur, installé dans une forteresse bien protégée. La forteresse est située au centre de la nation, entourée de villages et de fermes, et plusieurs routes y mènent. Un général décide de défier le dictateur et d'attaquer la forteresse. Il sait qu'en mobilisant d'un seul coup toute son armée, il a de bonnes chances de gagner. Il réunit donc toutes ses troupes à l'embouchure d'une des routes qui mènent à la forteresse mais, là, il apprend que le dictateur a parsemé les routes de mines afin que seulement peu d'hommes puissent y passer en même temps sans déclencher d'explosion. Un large contingent d'hommes qui passerait sur une route simultanément ne pourrait que détoner les mines. Le contingent aussi bien que les villages qui entourent les routes seraient ainsi détruits. Le général échafaude un plan : il divise ses troupes en petits groupes et les lance chacun sur une route différente. Les hommes se retrouveront tous devant la forteresse, mais sans passer par une unique voie. Il lance l'attaque et il en sort vainqueur.

Nous avons cette histoire à disposition. Nous savons également que les deux histoires sont liées, et que la solution de l'une s'applique à l'autre. En effet, le problème est, au fond (au niveau de sa structure profonde), le même. Dans cette condition, le pourcentage de participants à l'étude qui arrivent à la bonne réponse pour le problème de la radiation de Dunker monte à 92 % (Gick & Holyoak 1980, 1983). Deux éléments sont utilisés : un exemple analogue et l'explicitation de la nature commune du problème et de la solution. Voici comment les choses se sont passées dans le cadre de l'expérience de Gick & Holyoak (1980). Premièrement, ils ont demandé aux participants de lire l'histoire du général ainsi que deux autres histoires (des distractions n'ayant rien à voir avec le problème à résoudre). Les participants devaient mémoriser les trois histoires et en conserver en mémoire les éléments clés. De cette manière, les auteurs garantissaient que les éventuelles difficultés de résolution du problème posé ensuite n'étaient pas à attribuer à un défaut de mémoire du problème du général. Ils ont divisé les participants en deux groupes. L'un, le groupe expérimental, a reçu la consigne de résoudre le problème du chirurgien. Il a également reçu la suggestion explicite d'utiliser l'une des histoires apprises pour s'aider à résoudre le nouveau problème. Le groupe témoin n'a lui pas reçu cette suggestion. Onze sur les douze sujets du groupe expérimental ont utilisé la « bonne » solution, correspondant au problème du général. Dix de ces onze ont déclaré que l'exemple du général leur avait été très utile pour résoudre le nouveau problème.

Dans le groupe témoin, qui avait pourtant lu l'histoire du général et l'avait mémorisée, trois sujets sur quinze ont atteint la « bonne » solution – donc environ 20 %. La conclusion tirée par les auteurs considère que des analogies avec des domaines lointains peuvent aider dans la résolution de problèmes mais que percevoir cette analogie et l'utiliser n'est pas chose facile. Être exposé à un exemple analogue aide en effet, mais pas autant que y être exposé tout en sachant qu'il faut chercher une analogie utile dans l'exemple en question. En réalité, le rôle de l'explicitation semble crucial dans le cas présenté. Pourquoi, en l'absence d'une suggestion explicite, les sujets tendent-ils à ne pas percevoir l'analogie utile ?

4.4.1.2 Similitudes et analogies : mais à quel niveau ?

Le concept du niveau d'abstraction est clé dans la littérature sur le transfert par analogie. Revenons à l'exemple du général et du médecin. Il existe plusieurs similitudes entre les deux récits. Par exemple, dans les deux cas, le problème dispose d'une position géographiquement centrale. À un niveau plus profond – en réalité plus abstrait d'analogie –, dans les deux cas, il s'agit de vaincre un ennemi placé au milieu d'un environnement à préserver. Or, si nous nous représentons les deux histoires sans aucun détail spécifique, en faisant abstraction de la position de la forteresse et de celle de la tumeur, nous risquons de rater un indice d'analogie pouvant nous amener à transférer la solution d'un problème à un autre. Mais, si nous nous noyons dans les détails liés au domaine spécifique à chaque exemple, nous risquons dans ce cas de rater les similitudes. En effet, nous restons, par exemple, au contenu médical de l'un et pas de l'autre. Il semble donc raisonnable de se poser la question du niveau optimal d'abstraction qui facilite la reconnaissance d'éléments similaires entre les deux histoires – utiles pour la solution – et rend donc possible son utilisation. Comment les analogies sont-elles remarquées ? Comment arrivons-nous à connecter des informations, qui appartiennent à des corpus de connaissances et à des domaines sémantiques éloignés l'un de l'autre ? Pour un enseignant, ce type de question est un problème bien connu : comment faire en sorte que l'élève qui a traité une connaissance ou compétence dans un cours disciplinaire comprenne que la même connaissance ou compétence peut se représenter dans un autre cours, peut-être dans le cadre d'une autre discipline ou du même sujet ? Le problème du transfert analogique qui semble crucial pour la résolution de problèmes l'est donc aussi de façon générale pour l'apprentissage scolaire et pour l'esprit critique. En effet, une compétence de l'esprit critique apprise à l'école n'a de sens pour l'élève que s'il est capable de la mobiliser dans le contexte de sa vie quotidienne : au contact avec des informations, le « snapchat » d'un ami, un constat ou une intuition qui lui appartiennent.

Nous voyons donc que le problème du transfert et celui de l'analogie sont fortement liés[244]. Dans la littérature sur l'analogie, ce type de question se traduit en termes d'évocation et de *mapping*. Évocation et *mapping* sont considérées comme deux étapes fondamentales dans la formulation d'une analogie. L'évocation est le processus pendant lequel une histoire, formulation, texte… en évoque un autre. Dans l'exemple du général et du chirurgien, il s'agit de comprendre quels éléments d'une histoire permettent de faire penser à l'autre comme étant significative pour le problème à résoudre. Le *mapping* consiste ensuite à mettre côte à côte les éléments semblables pour réussir à voir en quoi la solution de l'un peut être transformée et traduite en une solution pour l'autre. De plus, évocation et *mapping* ne seraient pas aidés par le même type d'éléments et de ressemblances. Ainsi, du moins selon une partie des auteurs, l'évocation serait plutôt aidée par des similitudes superficielles, ayant trait aux contenus sémantiques. En revanche, le *mapping* serait plutôt soutenu par des similitudes dans la structure des deux (ou plusieurs) blocs à comparer. À l'inverse, les limites de l'analogie seraient liées à celles du contexte superficiel et à l'attraction qu'il exerce sur notre attention et mémoire. Pris dans les éléments de surface, nous ne parviendrions pas à percevoir ce qui se trouve plus en profondeur. Ainsi, si les cas du général et du chirurgien sont bien analogues en profondeur, le contexte de surface est en revanche tellement différent que l'analogie profonde n'est pas perçue parce qu'elle n'attire pas également l'attention.

Gick & Holyoak (1980, 1983) et Holyoak & Koh (1987) ont ainsi trouvé que, dans cet exemple, le recours au premier problème résolu pour résoudre le deuxième (le transfert) a lieu spontanément si les deux problèmes présentent des similarités au niveau de la surface autant qu'au niveau de la structure profonde, que le transfert n'est pas spontané mais a lieu si les deux problèmes diffèrent à la surface et ont des similitudes structurelles, si les sujets sont informés que le premier problème est utile pour le deuxième. En pratique, le transfert spontané a lieu mais seulement dans des conditions particulières ; si celles-ci n'existent pas, le transfert n'a lieu que si certaines similitudes moins intuitives sont explicitées. Nous retrouverons par la suite ces indications propres à la littérature sur l'analogie, sous forme de stratégies pratiques pour favoriser le transfert des apprentissages.

Dans la littérature sur l'analogie comme dans celle sur le transfert sont distingués les éléments profonds (qui appartiennent à la structure du problème) et les éléments de surface (qui devraient être ignorés). Nous remarquons par exemple que les experts dans un domaine sont moins distraits par la surface du problème et ont plus de facilité à percevoir et à utiliser le niveau de la structure. Par exemple, si deux problèmes se ressemblent superficiellement, mais pas au niveau de la structure, les experts n'en utilisent pas un pour résoudre l'autre (Novick 1988). Ils se situent à un niveau d'abstraction qui n'est pas celui du novice, notamment parce qu'ils possèdent des connaissances formulées de manière abstraite (des lois, des principes, des connaissances généralisables).

Reste cependant la difficulté de définir et de distinguer clairement les similitudes de surface et les similitudes profondes. De fait, nous pouvons à chaque fois identifier différents types de similitudes qui concernent les objets représentés dans les deux problèmes ou histoires, les relations logiques avec d'autres objets, les actions performées, les événements ayant lieu. Mais les sujets peuvent interpréter ces propriétés de manière différente en raison de leurs connaissances et expériences préalables, par exemple (Bassok & Olseth 1995). Par conséquent, ce qui compte comme une similitude peut être clair aux yeux des experts, qui se situent au bon niveau d'abstraction lorsqu'ils interprètent le problème, mais pas pour des novices, qui encodent les situations sur la base de leurs connaissances limitées du domaine. Un encodage à un niveau trop spécifique, trop détaillé, peut également rendre difficile la perception de l'analogie utile. Si le problème se situe au niveau de l'encodage – de l'interprétation –, alors une stratégie pour résoudre des problèmes en utilisant les bonnes analogies consiste à encoder correctement le problème – au bon niveau d'abstraction – en ignorant les détails trop spécifiques qui appartiennent à la « surface » du problème (Gamo, Sander & Richard 2010 ; Raynal, Clement & Sander 2017).
En conclusion, le recours à l'analogie est un outil puissant pour le transfert car la première est, au fond, à la base du second. Cependant, le transfert rencontre des obstacles spécifiquement liés à la capacité d'utiliser les « bonnes » analogies. Des stratégies doivent alors être mises en place pour aider à percevoir et à utiliser celles-ci. Ces stratégies peuvent être situées au niveau de la formulation des situations problème (multiplier les similitudes même de surface pour encourager à transférer) ; elles peuvent consister dans l'explicitation de similitudes au niveau profond, utiles pour la résolution ; ou elles peuvent consister en des aides afin d'encoder la situation à un niveau plus abstrait et général.

Imaginons vouloir aider un élève à reconnaître certains biais de raisonnement communs à plusieurs situations et contenus différents, comme le biais de confirmation. L'enseignant pourrait lui montrer des exemples où les ressemblances de surfaces sont suffisantes pour reconnaître l'analogie entre les exemples. Ensuite, il pourrait proposer à l'élève de résoudre un problème comportant un biais préalablement rencontré et résolu, et expliciter le fait que le nouveau problème est au fond analogue au précédent. Enfin, il pourrait aider l'élève à encoder chaque situation présentée à un niveau plus abstrait, notamment en expliquant les biais de raisonnement, et proposer une formulation générale de chaque biais (biais de confirmation, biais de sélection, etc.) afin d'aider l'élève à les reconnaître dans différentes situations.

4.4.1.3 Limites de l'analogie

Si l'analogie est un moyen puissant pour transférer connaissances et capacités d'un contenu ancien à un nouveau, cela comporte des limites et des risques. L'analogie peut en effet entraîner dans une direction indésirable et évoquer une connaissance ou compétence inappropriée à la situation, par la voie de similitudes de surface. Le cas des conceptions naïves ou préconceptions, qui constituent un obstacle aux nouveaux apprentissages, est bien connu. Par exemple, beaucoup d'étudiants ont du mal à retenir l'idée selon laquelle les différences entre les saisons sur Terre ne dépendent pas réellement de la distance Terre-Soleil mais de l'inclinaison de l'axe terrestre. Joue probablement dans cette méconception résistante le constat fait par l'expérience que la distance avec un corps chaud influence grandement la chaleur ressentie. Une connaissance expérientielle est donc mobilisée pour expliquer un phénomène qui se trouve à une échelle inaccessible, par analogie ; mais l'analogie est seulement superficielle et ne prend pas en compte des différences significatives pour le phénomène à expliquer. D'autres expériences montrent que le fait d'être exposés à un certain type de problèmes – et à leur solution – peut influencer négativement l'efficacité du raisonnement dans des problèmes superficiellement semblables mais qui requièrent un type de solution différente.

Imaginons que nous ayons à résoudre le problème suivant :

Nous devons mesurer une certaine quantité d'eau et nous avons à notre disposition plusieurs carafes de dimensions différentes. Par exemple, nous devons mesurer 100 ml d'eau et nous disposons d'une carafe qui peut contenir 21 ml, une carafe de 127 ml et une de 3 ml. Que faire ? Puis, nous devons mesurer 99 ml avec 14 ml, 63 ml, 125 ml. Puis 5 ml avec 18 ml, 43 ml, 10 ml. Puis 12 ml avec 9 ml, 42 ml, 6 ml. Puis 31 ml avec 20 ml, 59 ml, 4 ml. Puis 20 ml, avec 23 ml, 49 ml, 3 ml. Puis 18 ml avec 15 ml, 39 ml, 3 ml. Enfin, nous devons mesurer 25 ml avec des carafes de 28 ml, 76 ml et 3 ml.

Les personnes qui résolvent tous les problèmes en série ont généralement des difficultés avec le dernier. Celui-ci ne pose pourtant pas de problèmes s'il est présenté de manière isolée. Cela signifie que les solutions données au préalable influencent notre capacité à chercher une solution dans une direction nouvelle et qu'elles peuvent nous empêcher de trouver la bonne solution si nous sommes ancrés par analogie sur une solution inappropriée (Luchins & Luchins 1970).

Ce problème avec les analogies suggère que celles-ci doivent être dûment contrôlées. Toutes ne sont en effet pas productives. Le problème du transfert négatif nous alerte aussi sur l'importance d'évaluer correctement les stratégies que les élèves mettent en place dans la résolution de problèmes. Quand ils butent ou qu'ils retombent sur une solution erronée, la cause peut être l'utilisation de la mauvaise analogie, appelée par les circonstances ou particulièrement présente à l'esprit de l'élève à ce moment-là.

4.4.2 Avantages de la variation des contextes

Le contexte agit à plusieurs niveaux. Il existe de nombreux exemples d'apprentissages « de la rue » qui ne sont pas transférés dans des situations plus académiques, de classe, et, inversement, des apprentissages de classe qui ne sont pas re-exploités dans la vie de tous les jours. Ainsi, des enfants peuvent être capables de réussir certains calculs dans le cadre de la vente d'objets dans la rue sans être ensuite capables de répondre à des problèmes analogues en classe (Carraher 1986 ; Bransford, Brown & Cocking 2000). Le fait que le contexte semble peser négativement sur le transfert induit à penser que les apprentissages concrets, ancrés sur des cas réels, sont potentiellement négatifs. En effet, des apprentissages ancrés sur des problèmes réels à résoudre sont plus réfractaires au transfert dans de nouvelles situations. Pourtant, ceux-ci semblent en même temps faciliter compréhension et apprentissage – par

exemple, dans des domaines comme la médecine, apprendre à partir de cas concrets et de résolution de problèmes semble une méthode efficace. Comment résoudre cet apparent conflit entre compréhension et réutilisation ? Une manière, aussi présente au fond dans le cas du général et du chirurgien, consiste à multiplier les contextes – partir d'un cas spécifique, puis proposer des cas semblables – et demander de comparer ces contextes. Une autre consiste à demander de produire des cas semblables. Minervino, Olguin & Trench. (2017) ont proposé à leurs sujets le cas du général et d'autres cas analogues et ils ont demandé à l'un de leurs groupes de produire un cas semblable. Puis, ils ont placé tous les groupes devant le problème du chirurgien. Ils ont montré que la production influence positivement la capacité à se souvenir d'analogies profondes dans le temps. Si le transfert vers la vie quotidienne est celui visé, il faut aussi inclure, parmi les contextes présentés, des contextes de la vie quotidienne vers lesquels nous espérons voir s'effectuer le transfert[245]. Ceci inclut des contenus vers lesquels nous voulons que le transfert se fasse, mais aussi des manières de faire et de travailler – par exemple, le fait de savoir collaborer avec d'autres personnes.

4.4.3 Avantages de l'explicitation

Le cas du général et du chirurgien indique également que varier les contextes ne suffit pas ou n'est pas nécessairement efficace si l'apprenant n'est pas capable de percevoir qu'il existe des analogies entre ces contextes, à un niveau profond. Il s'agit donc de pointer, de mettre en évidence, de rendre perceptibles ces analogies. Ceci peut passer par différentes formes d'explicitation. Par exemple, dans le cas de l'étude menée par Gick et Holyoak (1981, 1983), nous pourrions faire remarquer qu'il existe une structure commune ou rendre celle-ci encore plus explicite en formulant le principe général qui sous-tend les deux cas. Donc en ayant recours à un enseignement direct, explicite du concept général à retenir. Un exemple historique illustre cette seconde manière de faire et ses effets positifs. Des enfants doivent centrer une cible avec des fléchettes ; la cible est couverte par quelques centimètres d'eau, et sa position visuelle est donc déplacée par la réfraction. Un groupe d'enfants reçoit un petit cours concernant la réfraction ; l'autre non. Les deux s'entraînent sur une même cible et réussissent autant. Mais, quand la quantité d'eau, change le groupe qui a reçu une explication explicite à propos du principe de la réfraction réussit mieux que l'autre (Scholkow & Judd 1908, in Bransford, Brown & Cocking 2000). Il s'agit donc, dans ce cas, de mélanger apprentissages ancrés sur des cas concrets et instruction amenant à se représenter le problème à un niveau plus abstrait (Singley & Anderson 1989 ; Bransford, Brown & Cocking 2000).

Dans certains cas, l'instruction explicite est fonctionnelle parce qu'elle permet d'isoler un critère discriminant qui s'applique au plus grand nombre de cas. Un cas bien connu est celui de la distinction du sexe des jeunes poussins. Cette tâche, perceptive, est particulièrement difficile pour des novices, mais s'améliore fortement avec l'expertise : les experts sont capables de distinguer le sexe, mâle ou femelle, d'un millier d'individus en une heure avec une fidélité de 98 %. Cependant, comme la cloaque des nouveaux nés est semblable pour les mâles et pour les femelles, les novices ne savent pas exactement ce qu'ils doivent regarder pour distinguer entre les deux : une région critique existe, plus ronde chez le mâle. Biederman & Shiffrar (1987) montrent que, si nous expliquons clairement cette différence à des sujets naïfs, leur performance s'améliore au point de ressembler tout de suite à celle des experts, comparée à un simple nouvel entraînement pour observer sans être instruits sur ce qu'il faut regarder. Fournir des critères que les experts utilisent (implicitement ou explicitement) permet donc de s'améliorer plus rapidement et de connecter plusieurs exemples entre eux. De façon générale, enseigner aux apprenants à représenter un problème à un niveau plus abstrait permet aux différents cas et exemples concrets présentés de cesser d'exister en tant que cas individuels et d'entrer dans une plus large famille ou schéma (Gick et Holyoak 1983 ; Bransford, Brown & Cocking 2000[246]).

4.4.4 Voie basse et voie haute

La variation des contextes, les aides ou supports à l'abstraction et le rôle des similitudes sont très présents dans l'une des approches classiques du transfert, celle de (Salomon & Perkins 1987). Ils ont distingué deux voies ou modalités propices au transfert.

D'une part, la voie basse passe par une importante pratique extensive et, dans une large variété de situations, jusqu'à l'automaticité. La généralisation est produite grâce à cette variété de situations sur lesquelles la capacité est pratiquée. Cependant, le transfert dépend fortement de la présence de similitudes de surface, perceptives. Si celles-ci ne sont pas assez fortes, elles doivent être appelées et rendues évidentes. Salomon et Perkins proposent l'exemple d'un apprentissage procédural par excellence : celui de la conduite. Avoir appris à conduire une voiture et l'avoir fait dans des conditions assez variées rend ensuite l'apprentissage de la conduite d'un camion plus facile : certaines compétences peuvent être transférées. Le nouveau contexte (celui de la conduite du camion) est en effet si perceptivement semblable au contexte original d'apprentissage qu'il appelle et déclenche presque automatiquement les schémas et routines de comportement appris dans le contexte de la voiture.

D'autre part, la voie haute passe par un exercice volontaire et réflexif. Il s'agit de faire abstraction du contexte pour extraire de celui-ci un principe d'ordre général. L'élément fondamental dans ce cas est l'abstraction avec explicitation d'un principe – la dé-contextualisation du principe à transférer. Celui-ci peut être utilisé de deux manières. D'un côté, l'apprenant peut extraire un principe abstrait et le garder à l'esprit en préparation de situations à venir où l'appliquer ; par exemple, il peut apprendre un principe de physique et se demander dans quels autres contextes il s'applique. De l'autre, l'apprenant peut se trouver dans une nouvelle situation présentant un problème à résoudre, en extraire des régularités et des caractéristiques, puis rechercher des correspondances avec des situations analogues déjà rencontrées. Dans les deux cas, il s'agit de faire abstraction du contexte et chercher des connections avec d'autres contextes et contenus. La voie haute ne dépend donc pas de l'existence de ressemblances de surface – à la limite celles-ci peuvent devenir confondantes – mais de la recherche des analogies dans la structure profonde des problèmes – ce qui n'est pas toujours facile.

Les deux modes peuvent être combinés. Il est possible de s'exercer dans une capacité jusqu'à l'automatiser (un calcul de probabilités par exemple), ensuite de réfléchir aux principes et à l'utilité de ce type de calcul à appliquer dans des situations différentes. Cependant, distinguer les deux modes est utile pour savoir dans quelles situations le transfert est possible ou difficile. Quand il passe par la voie automatique, basse, le transfert ne pose pas de problèmes majeurs. Il s'agit en tout cas, probablement, de transfert proche – à cause de la similarité des situations. Le transfert par la voie haute permet lui de relier des contextes plus éloignés, donc il ouvre la voie au transfert lointain, mais demande un effort de volonté pour abstraire et la capacité de créer des connexions. L'enseignant peut aider l'élève à transférer par la voie basse comme par la voie haute.

Premièrement, il s'agit de mettre à disposition de l'apprenant les situations que nous voulons connecter. Imaginons vouloir qu'un jour l'élève soit capable de réutiliser les connaissances apprises en histoire à l'analyse de faits politiques contemporains, ou celles apprises en biologie à un contexte écologique. Une possibilité pour l'aider à réaliser ce transfert consiste à lui présenter les faits contemporains et d'écologie dans le même contexte que celui dans lequel sont présentés les faits historiques et les connaissances en biologie. Il s'agit donc de créer des situations d'apprentissage qui contiennent ou sont semblables à celle désirée.

Deuxièmement, il s'agit d'aider l'élève à construire des liens entre ces situations, sans attendre qu'il le fasse lui-même et seul : par exemple, en lui fournissant les principes généraux et en lui explicitant en quoi ils sont pertinents pour le contexte, ou en

l'accompagnant à chercher lui-même ces généralisations (en posant des questions, en sollicitant des analogies, en invitant à appliquer le même principe à des domaines différents, à des matières différentes). En d'autres mots, il s'agit de faciliter l'abstraction.

En anglais, ces deux techniques sont appelées « *hugging* » (rapprocher) et « *bridging* » (faire le pont). Il s'agit pour l'enseignant de les utiliser de façon systématique et en sachant quand et pourquoi l'une et l'autre sont mises en place, pour faciliter quel type d'apprentissage et de transfert – proche ou lointain, par la voie basse ou haute. Salomon & Perkins (1987) suggèrent à l'enseignant de ne pas garder pour soi ces techniques, c'est-à-dire de les utiliser pour favoriser l'apprentissage et le transfert des élèves, et de les partager avec les élèves eux-mêmes. Convaincre les élèves qu'ils peuvent se demander eux-mêmes comment apprendre à transférer leurs apprentissages et à adopter des techniques pour s'auto-aider à transférer dans le futur : s'habituer à penser à des similitudes, par exemple, entre ce qu'ils font dans une matière et dans une autre ; entre ce qu'ils font à l'école et dans la vie quotidienne. Pour réussir cela, cependant, il semble nécessaire de penser toute l'éducation à la lumière du transfert. Par exemple, donner plus d'importance à des notions transdisciplinaires ou interdisciplinaires qui créent le contexte (par voie basse) pour transférer, ou accentuer des principes de base du raisonnement et ses limites qui sont inévitablement communs à toute matière et discipline : des tendances naturelles comme celle de rester coincé dans ses opinions, ne pas prendre en considération les hypothèses alternatives, etc. Rendre ces tendances explicites, faire abstraction du contexte et les relier à des cas concrets divers permet de développer les conditions pour un transfert de haut niveau, par voie haute, et de compétences assez générales pour investir tout type d'apprentissage. D'autres principes et règles se trouvent à mi-chemin entre les similarités de contenus et les abstractions de type métacognitif : le besoin de mesures et de preuves, de règles comme celles de la méthode, qui s'appliquent facilement à plusieurs domaines notamment scientifiques mais pas uniquement[247].

4.4.5 Beaucoup de pratique, mais pas n'importe laquelle

Tim Van Gelder – philosophe et scientifique cognitif s'étant intéressé à l'EC – compare la pensée critique au ballet (Van Gelder 2005[248]). Selon lui, la pensée critique n'est pas une capacité ou un ensemble de capacités sélectionnées par l'évolution, car elles ne sont pas nécessaires pour la survie dans un environnement comme celui dans lequel ont évolué nos ancêtres. Nous serions plutôt des organismes portés à chercher des patterns et des narrations cohérentes, sans pour cela nous demander si le pattern est réel et l'histoire vraie. Nous pouvons certainement discuter si des capacités propres à l'esprit critique n'auraient pas une origine évolutive ancienne, comme celles qui nous amènent à rechercher des patterns et à écouter des histoires – car tous ces patterns et ces histoires n'ont pas, en réalité, le même attrait sur nous. Cependant, nous pouvons accepter en première approximation ce cadre simplifié, qui amène Van Gelder à proposer que l'esprit critique doive être volontairement, et avec effort, appris, comme nous apprenons le ballet, le tennis ou une deuxième langue. Van Gelder explique en effet que, comme ces capacités, même si l'esprit critique est composé de composantes plus élémentaires, un apprentissage est nécessaire pour les combiner correctement – et non seulement pour les acquérir individuellement. Pour cette raison, apprendre à penser de façon critique demande du temps et de la pratique dédiée. Il s'inspire donc du domaine de l'expertise, notamment de l'approche développée par Ericsson et nommée « pratique délibérée » (Ericsson, Krampe & Tesch-Römer 1993). Selon cette approche, n'importe qui peut développer ses performances dans un domaine et en devenir expert, moyennant une longue (5 000 à 10 000 heures au total, 10 ans pour 4 heures de pratique quotidienne) pratique qui présente de plus ces caractéristiques : elle est menée avec concentration et dans le but de s'améliorer ; elle utilise des exercices spécifiques pour améliorer ses performances ; ses exercices deviennent progressivement plus compliqués pour monter en niveau ; elle est guidée et évaluée, couronnée par un *feed-back*[249].

Van Gelder reconnaît que le transfert reste un problème crucial pour la pensée critique – à la fois parce que, sans transfert, cet apprentissage perd son sens et parce que le transfert de cet ensemble de capacités apparaît particulièrement difficile et incertain[250]. La première étape consiste donc à reconnaître le problème, la deuxième à enseigner le transfert sans s'attendre à ce que celui-ci se fasse tout seul.

4.4.6 La métacognition et les apprentissages métacognitifs

Le transfert peut être décrit comme la capacité à mettre en pratique une connaissance dans un nouveau domaine ou contexte par rapport à celui d'apprentissage. Dans le cadre de cette définition, le transfert est un phénomène discret, qui a lieu ou n'a pas lieu. Nous avons cependant vu qu'il n'est pas toujours facile sans aide – et notamment sans la suggestion de regarder dans une certaine direction et chercher certaines similitudes. Ceci indique que le transfert peut être un fait de degrés. Nous pouvons, par exemple, avoir besoin de plus ou de moins d'aide pour transférer nos connaissances dans un nouveau domaine (aides génériques ou spécifiques, limitées ou nombreuses, etc.). Nous pouvons aussi considérer le transfert comme la capacité à apprendre dans un nouveau domaine, en nous appuyant sur des apprentissages effectués par ailleurs et préalablement (Bransford, Brown & Cocking 2000). Ces considérations se traduisent en des modalités d'évaluation du transfert, par exemple grâce à des mesures basées sur le nombre et le type de « coups de pouce » qu'il est nécessaire de fournir pour obtenir un transfert.

Il semblerait cependant que joindre des approches métacognitives à l'instruction aide le transfert, même en l'absence de « coups de pouce ». Ces approches métacognitives varient beaucoup dans leur format mais elles ont en commun l'idée de porter l'attention de l'apprenant sur les stratégies qu'il adopte pendant son apprentissage, les progrès qu'il réalise et les ressources qu'il mobilise (Bransford, Brown & Cocking 2000). Des exemples d'approches métacognitives sont en réalité des exemples d'apprentissage en enseignant ou enseignement réciproque, où les élèves acquièrent des connaissances afin de les transmettre à d'autres. Dans le processus, les apprenants sont amenés à réfléchir à ce qui facilite l'apprentissage et la compréhension et à évaluer leurs stratégies (Brown, Palincsar & Armbruster 1984).

Une considération relative à l'efficacité des méthodes d'enseignement doit être ajoutée à ces éléments. La pratique de la métacognition en classe (pratiques réflexives sur ce qui a été appris, explicitation, évaluation de l'apprentissage, apprentissage des stratégies métacognitives pour mieux apprendre) a récemment reçu beaucoup d'attention, notamment suite aux évaluations positives réalisées par John Hattie (Hattie, Biggs & Purdie 1996, Hattie & Donoghue 2016 ; https://visible-learning.org/hattie-ranking-influences-effect-sizes-learning-achievement/) et par le *Education Endowment Fund* (EEF, https://educationendowmentfoundation.org.uk/tools/guidance-reports/metacognition-and-self-regulated-learning/). Comme l'instruction explicite, l'enseignement des stratégies d'apprentissage et des stratégies métacognitives est une partie importante – voire dominante – dans le panorama éducatif actuel. Nous pouvons souligner que le rapport du EEF concernant les stratégies métacognitives en classe considère ces dernières comme fortement connectées au contenu d'apprentissage[251]. Introduire des pratiques explicites, réflexives, d'auto-évaluation pourrait également aider l'EC : comment le savons-nous ? Sur quelles preuves nous nous basons-nous ? Connaissons-nous bien ce domaine, au point de nous sentir confiants ? Avons-nous raison d'être si affirmatifs ? S'agit-il d'une situation présentant des facteurs de risque de nous tromper (investissement émotionnel, manque de connaissances, etc.) ?

4.4.7 L'argumentation comme support de l'esprit critique social et comme outil pédagogique

La notion d'argumentation (ou de bonne argumentation) est souvent associée à celle d'EC. Pour certains auteurs, il est important d'adopter des méthodes argumentatives pour améliorer les capacités d'argumentation mais aussi celles de pensée, notamment à partir de l'âge du collège.

4.4.7.1 L'approche argumentative (1)

Deanna Kuhn défend l'idée selon laquelle insérer l'argumentation dans les programmes scolaires est fondamental, au même titre qu'un cours de mathématiques ou d'histoire. Elle considère l'argumentation comme une capacité générale qui, à son tour, inclut différentes formes de raisonnement. La nature non disciplinaire, non liée à un contenu spécifique de l'argumentation, rend plus difficile son enseignement mais Kuhn ne considère pas cela comme une objection. Elle a en effet développé une méthode d'enseignement de l'argumentation qu'elle considère efficace pour des élèves d'âge correspondant à celui du collège. La méthode, basée sur un entraînement au débat, à l'analyse d'arguments et à la production, aurait, selon Kuhn, un effet positif sur l'argumentation écrite au niveau individuel, sur la capacité d'un groupe à mener des débats tout comme, ce qui nous intéresse ici, sur la capacité d'analyser les arguments, les preuves à l'appui et donc d'évaluer une thèse (Kuhn, Hemberger & Khait 2015). L'approche proposée est de type implicite ; la méthode est donc basée sur l'exercice du dialogue et non sur l'instruction explicite (Kuhn, Hemberger & Khait 2015[252]). Cependant, les transcriptions des dialogues servent pour le *débriefing*, et des activités réflexives sont proposées afin d'analyser comment les débats se sont déroulés et ce qu'il faut en apprendre.

La méthode ne prend pas en compte la vérité d'un argument, mais plutôt des capacités rhétoriques et d'ouverture aux positions d'autrui : l'élève doit initialement apprendre à prêter attention aux positions des autres et chercher à répondre avec des contre-arguments, le but étant d'affaiblir les positions de l'adversaire. Seulement, une fois que cela est acquis, l'élève apprend à utiliser des preuves pour renforcer ou affaiblir des positions. De fait, l'élève pourrait ainsi apprendre à renforcer ses arguments en sélectionnant des preuves à l'appui et en ignorant celles contre sa position. Son adversaire doit réaliser le travail de contre-argumentation. Dans une optique d'éducation de l'esprit critique, celui-ci ne serait pas un objectif individuel, mais uniquement pour un groupe de personnes en train de dialoguer, voire de débattre à partir de positions opposées (Kuhn, Hemberger & Khait 2015[253]). Les élèves débattent entre eux par l'intermédiaire d'un logiciel présentant des questions allant de questions portant sur le quotidien à des questions éthiques liées à des décisions politiques ou de santé. Ils travaillent à deux pour élaborer un argumentaire commun « contre » deux autres élèves. L'intervention en son ensemble se base sur des phases de débat argumenté en petits groupes (avec recherche d'arguments et de contre-arguments), suivies d'un débat en grand groupe, puis d'un débriefing. Le cycle recommence à la fin de chaque intervention (treize semaines au rythme de deux rencontres par semaine). Certaines classes ayant participé à l'expérience ont travaillé deux ans sur le projet, d'autres trois ans. Le projet est mené par un intervenant externe. Des classes de contrôle suivent un programme de même intensité mais plutôt basé sur l'approche de Lipman de la philosophie pour enfants – sans dialogue structuré, travail en petits groupes et préparation de contre-arguments – et sont encadrés par un enseignant de l'école. La mesure d'impact est qualitative, micro-génétique, basée sur l'analyse des productions.
Au-delà de la difficulté de se prononcer sur les résultats obtenus par ce programme, lourd dans son implémentation, et bien que l'objectif soit d'améliorer la capacité de discuter à partir d'arguments solides, ce type d'intervention ne nous semble pas répondre à la définition d'EC proposée dans ce document. En effet, les élèves choisissent leurs arguments à partir de leurs

connaissances, probablement partielles. De plus, ils ne sont pas amenés à lire avant de débattre afin de s'informer (pour éviter de brider le débat ou de fournir des informations qui n'apparaitraient pas immédiatement pertinentes et intéressantes). Les élèves identifient eux-mêmes les questions qui demandent plus d'information, et l'enseignant leur fournit lors d'une séance plus avancée des informations qu'il sélectionne en réponse à ces questions. À aucun moment les élèves ne sont confrontés au besoin d'évaluer la qualité de l'information, son poids en termes de preuves à l'appui et de fiabilité des sources (Kuhn, Hemberger & Khait 2015[254]).

Kuhn a aussi proposé d'autres approches méthodologiques de l'argumentation, qu'elle met directement en relation avec l'EC. Ainsi, par exemple, Kuhn (2015) a exploré l'impact de l'argumentation sur les capacités d'EC d'élèves de collège (en réalité, l'impact d'un groupe de leçons explicites portant sur l'analyse des arguments). Le cours est considéré comme un cours de philosophie et se compose de huit séances de trente minutes, au rythme de deux rencontres par semaine, pendant lesquelles un intervenant explique de façon explicite la valeur d'écouter les arguments des autres (*open-mindedness*), comment construire des contre-arguments, l'importance de se questionner à propos de ses croyances, d'analyser la force des arguments sur la base de leur construction et des preuves à l'appui. Les séances comprennent également des moments de débat et de discussion entre les élèves ayant des vues opposées afin d'analyser de façon collaborative les contre-arguments des uns et des autres. La mesure d'impact consiste en un pré-test/post-test composé de six questions ouvertes concernant l'analyse d'arguments. Les questions sont proposées sous forme d'entretien. Les élèves sont amenés à justifier leur opinion concernant une option qui ne demande pas de connaissances particulières pour se prononcer (par exemple, la durée de la journée scolaire). Les évaluateurs évaluent (qualitativement) la capacité de l'élève à donner une opinion, à apporter des raisons à l'appui ou à expliquer comment il ferait pour chercher des preuves à l'appui (quel type de recherches seraient nécessaires, quelles sources pourraient-être interroger, comme établir leur fiabilité…), à prendre en compte les options opposées et les arguments en faveur de celles-ci. D'autres questions évaluent la métacognition, la réflexion et la signification que les élèves donnent à l'idée de « mieux penser ».

Dans seule une école, 24 élèves volontaires ont participé à l'étude (26 moins deux qui n'ont pas répondu aux questions). Ils ont été divisés en trois groupes, dont deux de traitement et un de contrôle. L'étude présente un résultat positif du cours de philosophie, en montrant une évolution de la qualité des réponses entre pré-test et post-test uniquement pour les élèves des groupes de traitement. Cependant, ce résultat est en réalité de faible valeur de preuve vu le pauvre nombre de participants et la nature qualitative de la mesure. Aucune mesure n'est fournie dans le temps, ni dans un cadre « écologique » différent de celui auxquels les élèves ont été entraînés en classe pendant le séminaire. En particulier, le format du test correspond à celui de la dernière journée de séminaire, quand les élèves sont amenés à débattre entre eux de l'existence du Père Noël.

4.4.7.2 L'approche argumentative (2)

Hugo Mercier propose également une approche argumentative du raisonnement et de la résolution de tâches intellectuelles. Son approche est basée sur une vision de la raison comme étant, de par sa nature, argumentative : les capacités relatives à la « raison » (la recherche et la présentation d'arguments à l'appui d'une position) et au raisonnement (la recherche de solutions à un problème intellectuel, grâce aux connaissances disponibles) se seraient développées en réponse à un besoin de type social : celui de faire valoir ses arguments, de persuader. Lorsque nous cherchons des raisons, nous ne cherchons pas à prendre la meilleure décision, mais à convaincre les autres que nous sommes corrects. Les conséquences visibles du fait que la fonction propre à l'argumentation serait la persuasion et non la prévention des erreurs seraient que: a. notre raisonnement individuel et solitaire nous induirait souvent en erreur (Mercier & Sperber 2011 ; Mercier 2016 ; Boku, Yama & Mercier 2018[255]); b. la

situation de groupe, argumentative, donnerait de meilleurs résultats en termes de résolution de tâches intellectuelles (Boku, Yama & Mercier 2018[256]).

Quant au transfert des performances, la littérature sur l'argumentation n'est pas claire. Elle montrerait que le fait de participer à une situation de discussion argumentative augmente les chances de résoudre des problèmes analogues dans le futur (immédiatement après, dans la phase de post-test qui suit). Dans le cas des expériences décrites par Mercier et ses collègues (par exemple, Trouche, Sander & Mercier 2014 ; Boku, Yama & Mercier 2018), il s'agit de transfert proche : le transfert se fait sur une même structure du problème, avec un contenu différent. Cependant, le transfert ne fait pas partie en soi des prédictions de la théorie, car les expériences ne sont pas censées tester une approche éducative mais une situation de résolution d'un problème de raisonnement. En effet, le transfert proche permet de vérifier que les participants ont compris pourquoi la bonne solution au problème est celle donnée à la fin du débat, mais pas s'ils savent utiliser des stratégies de raisonnement différentes dans différents problèmes grâce au fait d'avoir pris parti à l'expérience. Il se peut donc que la participation à des activités de débat ou discussion argumentée ne serve qu'à atteindre une meilleure performance dans le cas spécifique, mais pas à apprendre une capacité transférable (Boku Yama & Mercier 2018[257]). Les tests ne comportent pas de groupe contrôle qui aurait à sa disposition le même temps que le groupe de traitement pour résoudre le problème (les participants cherchent à résoudre le problème individuellement pendant cinq minutes, et ceci sert de groupe de contrôle pour la phase de groupe qui dure dix minutes, après laquelle les participants ont encore cinq minutes pour répondre à l'énigme) (Boku, Yama & Mercier 2018[258]).

4.4.7.2.1 Argumentation et évaluation des arguments d'autrui

Le recours à l'argumentation dans la palette de méthodes pour l'enseignement de l'EC est cependant à prendre au sérieux pour la raison suivante, développée par Mercier (2016) : l'évaluation des arguments à l'appui d'une position semble favorisée dans une situation de discussion plus que dans une situation individuelle. Dans une situation individuelle gagnerait une attitude « flemmarde » à confirmer ses propres positions[259]. En revanche, dans une situation sociale, l'individu serait porté non seulement à chercher de bons arguments en faveur de sa position mais aussi à évaluer ceux de l'adversaire de manière plus objective et exigeante qu'il ne le ferait avec soi-même. Demander à quelqu'un d'évaluer ses propres arguments serait donc moins efficace que lui demander d'évaluer les arguments d'autrui. (Il est cependant plus difficile, dans un cadre argumentatif, de mesurer la capacité d'évaluation des arguments que celle de production de bons arguments[260].) En outre, Mercier (2016) défend l'idée selon laquelle, dans la vie de tous les jours, un sujet moyen est relativement performant dans l'évaluation de la force des arguments apportés par autrui. Mercier (2016) cite notamment des études qui montrent que les interlocuteurs identifient des arguments faibles car basés sur des idées fausses argumentatives (argument d'autorité relatif à une source non experte et non nécessairement bienveillante) ou sur des erreurs logiques[261]. La situation de dialogue propre aux échanges argumentatifs permettrait donc de mieux exploiter ces capacités par rapport à une condition d'évaluation solitaire. Elle représenterait alors une condition particulièrement favorable à l'exercice – si non à l'éducation – de l'EC.

L'apprentissage coopératif semble d'ailleurs donner des résultats positifs en éducation et dans le développement des capacités cognitives des jeunes enfants (Slavin 2014 ; Sills, Rowse & Emerson 2016). Kyndt et al. (2013) ont réalisé une méta-analyse des études sur l'apprentissage coopératif depuis 1995 et ont confirmé des effets positifs sur les apprentissages et les attitudes.

4.5. Synthèse et considérations pratiques pour les pédagogies de l'EC

La psychologie contemporaine semble sceptique quant à la possibilité d'apprendre des heuristiques ou règles générales et de les appliquer indépendamment du contenu. Même les plus optimistes (comme Richard Nisbett) reconnaissent les limites de ce type de généralisation : nécessité de posséder déjà un système de règles intuitif à substituer avec celui plus sophistiqué, transparence du problème (qui laisse percevoir le type de règles à appliquer), limitation des systèmes des règles générales à peu de typologies (règles statistiques, causalité, certaines règles sociales). Cependant, il ne semble pas irrationnel d'entretenir l'espoir de développer certains savoir-faire cognitifs de façon générale, au-delà d'un contenu et d'un contexte donnés, ou du moins de les voir transférer dans plusieurs situations différentes, du moment où ces savoir-faire :

- s'appuient sur des stratégies que nous pouvons apprendre et ensuite pratiquer de façon soutenue dans différents contextes ;
- s'expriment de façon abstraite de manière à les retrouver plus facilement et à développer des raccourcis dans notre mémoire.

Cet espoir doit prendre en compte les différentes formes de transfert qui peuvent être plus ou moins éloignées, en raison de la similitude entre les situations visées. Nous ne devons pas ignorer le rôle joué par les connaissances à la fois dans la capacité de résoudre des problèmes ou de penser de manière efficace, et dans les contraintes que cela impose au transfert dans des situations nouvelles où nous manquerions des connaissances nécessaires.

L'apprentissage de capacités générales de pensée, comme la pensée critique, ne semble pas pouvoir être disjoint de l'acquisition de connaissances. Pour bien penser dans un nouveau domaine, il faut également se doter de connaissances de domaine de qualité et en quantité suffisante. Posséder une base de connaissances riche et variée est une condition nécessaire pour bien penser. Les heuristiques générales ne sont pas suffisantes.

Nous pouvons espérer transférer – du moins à un certain degré – les capacités acquises dans la pratique d'un domaine particulier à un autre, si cette acquisition a suivi des principes d'apprentissage pour le transfert. En effet, le transfert ne se réalise pas tout seul ni automatiquement. Là où il ressemble à un transfert automatique, deux conditions sont en réalité probablement réunies : d'un côté, la pratique extensive a permis l'automatisation de la capacité dans un contexte ; de l'autre côté, le nouveau contexte présente des éléments de ressemblance forts avec le contexte d'acquisition, des attracteurs perceptifs qui facilitent fortement le transfert et, nous pourrions dire, qu'ils l'appellent même « contre » volonté.

Dans les autres cas de transfert, des heuristiques générales ou des principes abstraits ont été produits volontairement et raccrochés à des contextes concrets, à des pratiques. Ils semblent nécessaires pour faciliter, voire rendre possible, le transfert. Mais nous ne pouvons pas espérer que abstractions et bases de connaissances ou situations concrètes d'application se rejoignent toutes seules, sans l'aide de formes appropriées d'apprentissage et éventuellement sans des aides ponctuelles et locales invitant à prendre en compte les unes en relation avec les autres (suggestions, aides externes, facilitations sous forme de contextes similaires même à la surface).

L'éducation de l'esprit critique est l'histoire d'une troisième voie : l'éducation n'a pas besoin de renoncer à des capacités générales, de se retrancher uniquement dans l'acquisition de connaissances spécifiques et locales de peur de « perdre du temps » inutilement ; elle n'a pas non plus à choisir de développer les capacités de pensée tout en renonçant à fournir aux élèves des bases de connaissance riches ; les principes du raisonnement, de la pensée critique, les

méthodes d'investigation ne peuvent pas être enseignés et appris de manière abstraite, ils ne doivent non plus être pratiqués sans être accompagnés de principes et heuristiques générales et d'une réflexion sur la manière de transférer ces heuristiques dans un nouveau contexte.
Nous sommes souvent, dans la vie quotidienne, dans le cas de figure suivant : nous ne pouvons pas espérer acquérir des connaissances approfondies dans tous les domaines que nous rencontrons, mais nous sommes amenés à prendre une position, à exprimer une opinion et à décider dans une direction ou dans une autre.

Il semble alors nécessaire d'adopter un principe de réalité. Il est improbable voire impossible d'imaginer que nous puissions être également capables de penser « de façon experte » dans n'importe quel domaine. L'objectif n'est pas la perfection, mais une amélioration des capacités de réflexion et leur application plus satisfaisante dans des contextes et face à des contenus d'intérêt personnel ou général.

Revenons à la notion de méthode faible et de résultat faible. En utilisant des principes heuristiques généraux dans un nouveau contexte, nous pouvons obtenir un transfert et réussir à donner une réponse suffisamment satisfaisante – par rapport à quelqu'un qui ne disposerait pas de ces heuristiques. Cependant, le résultat ne sera probablement pas aussi satisfaisant que si, en plus des principes généraux, nous pouvions également compter sur des connaissances de domaine. Il s'agira alors d'un compromis entre le coût d'acquérir, en plus de principes généraux, de nouvelles connaissances de domaine et le gain d'une réponse optimale à la nouvelle situation.

4.5.1 Conseil pratiques pour favoriser le transfert de l'EC

Des conseils pratiques émergent de cette synthèse pour l'enseignement de l'EC. Nous devons cependant souligner qu'ils se basent sur des éléments qualitatifs d'analyse de la littérature et qu'ils nécessitent donc d'être mis à l'épreuve et testés avec des méthodes appropriées permettant d'en mesurer l'efficacité.

1. *Enseigner les compétences de l'esprit critique et les stratégies/connaissances utiles pour les nourrir de façon assumée*

- Enseigner les compétences relatives à l'esprit critique de façon assumée : ne pas espérer qu'en les infusant dans les « programmes », elles seront comprises, généralisées et transférées.

2. *Enseigner pour le transfert*

- Enseigner les principes généraux et stratégies générales en relation avec des cas concrets pour montrer comment ils peuvent être mis en pratique.
- Faire pratiquer de façon répétée et variée une capacité jusqu'à l'automatisation.
- Répéter dans des contextes variés mais proches.
- Exemples disciplinaires multiples.
- Répéter dans des contextes où nous voulons voir le transfert avoir lieu et mettre en avant les similitudes.
- Exemples de la vie quotidienne.
- Faire abstraire et expliciter pour favoriser le transfert réflexif et éventuellement l'aider avec des suggestions, questions, etc. si cela ne vient pas tout seul.
- Faire réfléchir à ce que nous sommes en train de réaliser en mobilisant la métacognition.
- Utiliser l'analogie au bon niveau de profondeur et de pertinence.

3. *S'appuyer sur un esprit critique « social »*.

- Avoir recours à des situations de débat argumentatif, des activités de groupe.

4. *S'appuyer sur des aides externes matériels*

- Utiliser des outils pour apprendre à transférer mais aussi pour apprendre à apprendre à transférer : faire une liste de ceux que nous avons utilisé dans différentes situations.
- Expliciter les raisons qui poussent à utiliser ce type d'outils, inviter à se les approprier.

5. Avis pédagogiques : choix à faire et à ne pas faire dans l'éducation de l'EC

Les apports théoriques que nous avons mobilisés dans ce rapport permettent de dégager un ensemble d'avis pédagogiques concernant l'éducation de l'EC : des voies à tenter et d'autres que la théorie ou la pratique, ou les deux, amènent à rejeter. Nous commencerons par ces dernières.

5.1 À ne pas faire

5.1.1 Ne pas réduire l'éducation de l'EC à une éducation aux médias de l'information

Un article publié en 2012 présentait une corrélation entre le nombre de prix Nobel obtenus dans un pays et la consommation annuelle de chocolat de la population de ce même pays (Messerli, 2012 ; voir aussi Maurage et al. 2013). Ce type de graphique incite le lecteur à une inférence presque immédiate, le faisant glisser du constat d'une corrélation à un lien de causalité : nous imaginons, même quelques instants, que la consommation de chocolat est la cause de l'obtention des prix Nobel, peut-être parce que le chocolat contiendrait des molécules ayant des propriétés sur le développement de nos capacités intellectuelles... Cet article a été repris dans différents médias français et étrangers : les titres sont parfois prudents dans leur formulation (« Plus un pays mange de chocolat, plus il a de prix Nobel »), d'autres moins prudents (« Manger du chocolat pour avoir le Nobel »). Une telle anecdote illustre le fait que nous n'avons nullement besoin des médias pour parvenir à ce type de conclusions. Les médias jouent un rôle indéniable en faisant circuler cette information mais l'information est par elle-même séduisante, au sens où elle possède des caractéristiques qui font qu'elle se propage rapidement. La première de ces caractéristiques est citée plus haut : nous inférons une relation de causalité sur la base d'indices dont la corrélation et cela peut nous amener à des illusions de causalité (Matute et al. 2015). Cette tendance prévaut d'une certaine façon dans l'existence de cette « fausse information ».

Nous avons souligné dans la troisième partie de ce rapport l'importance d'outiller les élèves de critères spécifiques et de connaissances adaptées au contexte, qui permettent une évaluation plus fine des sources d'information (par rapport aux critères et indices utilisés de manière spontanée). Une éducation de l'esprit critique qui ne se focaliserait que sur une éducation aux médias serait cependant insuffisante : dans l'exemple présenté ci-dessus, l'information a initialement été proposée par un médecin dans une revue prestigieuse. Se focaliser sur la source (solide dans ce cas) ou sur les preuves avancées (avoir suffisamment de connaissances en biologie pour discuter des faits sous-jacents à cette thèse est délicat… les élèves pourraient même arriver à se convaincre que la thèse est crédible) se révèlerait insuffisant car le cœur du problème est ailleurs. De plus, sans prendre plus de précautions, les élèves pourraient penser qu'il est nécessaire de se mettre à douter de chaque information, y compris celles délivrées dans des articles scientifiques. En réalité, l'auteur de l'article ne conclut pas à un lien de causalité, comme cela a été relayé dans certains médias.

La conclusion de cette réflexion n'est pas qu'il faut nier l'intérêt d'une éducation au monde des médias. Au contraire, les élèves doivent être sensibilisés aux caractéristiques des nouveaux médias. Cependant, cet apprentissage doit être relié à une réflexion sur ce qui, dans notre propre fonctionnement cognitif, peut nous amener en erreur. Par exemple, nous recherchons tous, spontanément, des causes aux phénomènes que nous observons, et repérer

des corrélations est un bon indice pour y parvenir. Mais le fait de constater une corrélation n'est pas suffisant : en dehors d'autres indices, cela reste un niveau de preuve encore faible. Comment pourrions-nous nous assurer d'un lien de cause à effet ? Un test expérimental ou la recherche de mécanismes causaux permettraient d'obtenir des preuves plus solides. Plutôt que de construire simplement une image sombre du monde des médias dans l'esprit des élèves, l'objectif est de leur donner des outils pour évaluer l'information qu'ils reçoivent.

5.1.2. Faire attention à ne pas donner l'impression que ce qui compte dans un débat argumentatif est de bien présenter ses arguments

Nous avons souligné dans la quatrième partie de ce rapport l'intérêt de mobiliser des situations d'apprentissage dans lesquelles les élèves sont amenés à débattre et à argumenter. Le raisonnement collectif pourrait favoriser leur capacité à repérer des mauvais arguments, ou des mauvaises structures argumentatives, alors qu'une même approche menée de manière solitaire les conduirait à se renfermer sur leurs positions. Il est cependant important de souligner que le travail sur l'argumentation ne peut pas se suffire à lui-même pour éduquer l'EC.

Un aspect à considérer quand nous travaillons l'argumentation est le risque d'un recours non encadré au débat. Certains enseignants pourraient décider d'organiser en classe des débats sur des questions vives de société (par exemple, le danger des vaccins ou des OGM) pour faire passer des messages généraux sur les attitudes à adopter : certains enseignants pourraient en profiter pour montrer qu'il est facile de mentir, de développer des arguments fallacieux, et inviteraient ainsi les élèves à « se méfier » ; de manière diamétralement opposée, d'autres pourraient déclarer qu'il n'existe pas une « vérité unique » mais un ensemble de points de vue. L'enseignant inviterait alors les élèves à faire preuve d'ouverture d'esprit, à ne pas rester ancrés sur leurs positions, et finalement à se montrer ouverts aux idées contraires au prétexte qu'aucune connaissance n'est figée (surtout pas dans le domaine des sciences). Que pourrions-nous reprocher à de tels messages ?

Premièrement, de tels conseils généraux sont difficiles à mettre en pratique et à décliner dans des contextes bien précis (se méfier de qui ? Être exigeant de quelle manière ? Comment faire preuve d'ouverture d'esprit ?).

Deuxièmement, ces conseils peuvent donner lieu à des contresens. Se fermer à toutes les informations de manière aveugle ou au contraire considérer que tout savoir peut-être remis en question, c'est négliger le fait que toutes les opinions ne se valent pas. Parce qu'aucune thèse n'est complètement sûre, nous regroupons dans la même catégorie des savoirs établis de manière rigoureuse, vérifiés par des faisceaux de preuves, et des points de vue personnels non étayés. En demandant aux élèves de défendre tout à tour des positions « anti » et « pro-vaccin » par exemple, l'enseignant les inciterait à exercer leurs compétences argumentatives en niant la supériorité de certains arguments (ceux appuyés par des faits) sur d'autres. Rester ouvert à toutes les positions peut mener à une posture de doute, à un relativisme qui s'oppose à la pensée critique. En revanche, se montrer méfiant par défaut, par excès de vigilance, c'est là aussi refuser de reconnaître la supériorité de certaines informations. Cette posture invite à se replier sur ses propres convictions personnelles, sans discernement.

Enseigner le débat doit au contraire créer les conditions pour que les élèves comprennent que l'ensemble des points de vue sur un sujet ne se valent pas : tous ne sont pas acceptables, pas plus que tous méritent notre suspicion. La valeur épistémique relative des informations à notre disposition est ce qui permet de trancher. Le débat doit mener les élèves à départager des points de vue sur la base de faits, plutôt que simplement s'exercer à défendre à tout prix une position.

Il est enfin important de réaliser que, lors d'un débat, nous associons souvent des positions qui relèvent de la préférence à des considérations qui relèvent de faits, appuyés par des preuves. Accepter ou non que la société ait recours à une technologie donnée sous l'angle de considérations éthiques relève de préférences, de positions personnelles. Ces préférences se valent *a priori* et se discutent. L'innocuité d'un vaccin donné relève, en revanche, du fait scientifique et ne se discute pas sous l'angle des préférences personnelles, mais des preuves qui sous-tendent l'affirmation. Placer les élèves en position de débat se prépare et s'accompagne, non pour amener les élèves à échanger librement et de manière ouverte sur tous les sujets, mais pour leur apprendre à distinguer les préférences, les opinions, les connaissances, les faits et les différents niveaux de preuves.

5.1.3. Ne pas se limiter à des thèmes chauds

Nous pourrions être tentés de focaliser l'enseignement de l'EC autour de théories scientifiques d'enjeu majeur comme la théorie de l'évolution ou le réchauffement climatique. Un tel choix peut se révéler contre-productif.

Tout d'abord, ces thèmes de discussion comportent, en plus d'un engagement cognitif, un aspect d'appartenance à un certain groupe social. Il devient particulièrement difficile de se départir de ses idées dans un tel contexte (voir la deuxième partie de ce rapport). Ainsi, un élève qui adopterait une position allant à l'encontre d'une théorie scientifique pourrait le faire sur une base complètement différente que celle du raisonnement, des arguments et des faits. Dans ce cas, ramener inexorablement la discussion sur les faits ou la construction des théories scientifiques ne suffirait pas nécessairement, car d'autres éléments seraient alors en jeu.

Ensuite, un risque existe à chercher à défaire un « mythe », une fausse croyance : celui de renforcer cette position. Certains élèves pourraient ne pas être familiers avec une pseudo-thèse (comme celle de l'*intelligent design*) et, dans ce cas, un enseignement qui viserait à déconstruire la pseudo-thèse aurait d'abord pour conséquence d'exposer les élèves à ce contenu. Il est aussi possible que les élèves ainsi exposés oublient, après quelques temps, que l'enseignant a précisé qu'il s'agissait d'une fausse information, pour ne garder en mémoire que le contenu. Plus généralement, vouloir déconstruire des fausses théories est un exercice risqué qui peut produire l'effet opposé à celui espéré (Cook, 2017).

Enfin, concevoir un enseignement de l'esprit critique sous la forme d'un projet ponctuel ou sur un thème sensible risque de focaliser toute l'attention des élèves sur le contexte en question (voir la quatrième partie de ce rapport). Dans une telle situation, les outils présentés par l'enseignant sont alors très dépendants du contexte ou tellement associés à celui-ci que les élèves ont de grandes difficultés à le transférer à d'autres théories scientifiques, et plus encore à des situations de la vie quotidienne. Il n'est donc pas possible de considérer que l'enseignant a développé l'EC des élèves. L'enseignement de l'EC ne peut donc pas s'appuyer en tout et pour tout sur un enseignement autour de questions vives.

Cela signifie-t-il qu'il ne faut pas aborder des thèmes comme la théorie de l'évolution ? Non, bien sûr. Au moment d'aborder des sujets plus délicats, l'enseignant peut remobiliser tous les outils de l'esprit scientifique et critique qu'il a enseignés jusqu'alors. Il serait illusoire de penser que les élèves pourront facilement les importer et ignorer tous les obstacles de leur compréhension de la théorie de l'évolution. Pour plusieurs raisons :

- lorsque les élèves découvrent cette thématique, ils ont peu de connaissances solides. Il est impossible d'exercer son esprit critique en dehors de connaissances ;;
- le contexte très chargé de cette thématique rend particulièrement difficile le repérage de la structure profonde des obstacles rencontrés (par exemple, le rôle du hasard, le raisonnement par les conséquences…) ;

- chaque théorie a ses propres écueils, ses propres obstacles, que nous retrouvons difficilement ailleurs et qui freinent la compréhension globale tant qu'ils perdurent.

Malgré tout, il est raisonnable de penser que l'utilisation en amont d'outils transférables (aussi bien la connaissance de nos propres limites que la connaissance plus épistémologique du fonctionnement de la science, de la nature du fait scientifique par rapport à l'opinion, etc.) peut favorablement influencer l'adhésion à des théories scientifiques majeures.

5.1.4. Ne pas limiter l'enseignement de l'EC à une prise de connaissance des biais cognitifs, et éviter les listes de biais

Une vaste littérature grand public (mais aussi scientifique) relative à l'EC se focalise sur les biais cognitifs et les erreurs de jugement qui ne permettraient aux êtres humains que d'avoir qu'une rationalité limitée (voir la troisième partie de ce rapport). Sont ainsi popularisées des longues listes de biais qui soulignent à quel point nous raisonnons mal (voir, par exemple, Wikipedia. *List of cognitive biases*).

Nous avons montré, dans la troisième partie de ce rapport, l'intérêt de comprendre que certaines circonstances nous mettent plus que d'autres dans un risque d'erreur, et l'intérêt de savoir identifier les obstacles cognitifs à une évaluation juste de la qualité des informations disponibles et à la calibration de sa confiance en ces informations.
Cependant, des limites existent à une telle approche pédagogique. Tout d'abord, celle-ci peut véhiculer l'idée selon laquelle notre cognition nous servirait en réalité mal. Ce type de considérations globales, à l'instar de celle selon laquelle « toutes les informations sur Internet sont fausses », ne permet pas d'instaurer un discours juste sur lequel s'appuyer. Or, nous reconnaissons parfaitement bien des visages, des émotions, nous connaissons des faits que nous pouvons attester avec nos sens (la pomme tombe bien vers le sol, en accord avec la théorie de la gravité). Bref, nos outils naturels nous servent bien la plupart du temps. Il est vrai, cependant, que certaines théories scientifiques sont contre-intuitives, que notre approche des statistiques et des probabilités est limitée, et que nos raisonnements s'appuient sur des heuristiques pas toujours adaptées. L'effort pédagogique à mener ne consiste donc pas à donner à l'élève une liste pessimiste de toutes les catégories d'erreurs que nous commettrions à chaque instant mais plutôt de souligner que certains domaines disciplinaires ou certaines situations (c'est le cas d'un graphique présentant une corrélation) nous font mal juger la qualité d'une information et qu'il est important d'apprendre à mieux identifier et donc à anticiper ces situations ou du moins à baisser la confiance dans ses opinions lorsque les circonstances sont plus favorables à des erreurs d'évaluation.

Insistons sur le fait qu'il ne s'agit pas – dans le cadre d'une éducation de l'EC – d'amener l'élève à abaisser systématiquement sa confiance, au risque de simplement semer le doute dans son esprit. Il s'agit plutôt de l'amener à reconnaître les situations dans lesquelles il peut légitimement accorder sa confiance et celles où l'information est moins fiable, voire franchement suspecte. Fort de cette connaissance, l'élève peut ainsi augmenter sa vigilance (ou abaisser sa confiance) dans ces situations bien précises, et non de manière aveugle, à tout moment.

5.1.5. Ne pas se limiter à des enseignements implicites, non assumés de l'EC, ni se limiter à un enseignement abstrait et théorique de celui-ci

Nous pourrions espérer que l'enseignement de la méthode scientifique se suffise à lui-même pour rendre les élèves capables de reconnaître des informations étayées par des preuves solides et les distinguer d'anecdotes. Les élèves découvriraient les outils de la science en

classe et les convertiraient sous la forme d'outils guidant leurs choix des meilleures informations. Nous avons malheureusement souligné toutes les raisons qui nous poussent à penser qu'un tel transfert spontané n'est pas réalisé (voir la quatrième partie). Plusieurs auteurs rappellent la difficulté, pour l'apprenant, de faire abstraction du contexte dans lequel il a découvert une notion, pour la remobiliser dans une situation nouvelle. L'apprenant a plutôt tendance à rester aveugle aux structures profondes des problèmes et happé par leurs structures superficielles. Le pari du transfert est d'autant plus délicat que la distance entre les domaines (domaine scolaire et vie quotidienne) est grande.

Un effort spécifiquement dédié au transfert doit donc accompagner tout cours de sciences (ou d'autres disciplines) qui présenterait un outil d'évaluation de l'information. D'autant que, nous l'avons souligné, l'exercice de l'esprit critique est aussi dépendant de connaissances. L'éducation de l'EC ne pourrait donc pas se faire au détriment des enseignements disciplinaires existants.

Il ne s'agit donc pas de minimiser les enseignements de contenus fondamentaux, mais de trouver une place pour un discours explicite sur des aspects plus abstraits. Ainsi, chaque fois qu'il est en mesure de le faire, l'enseignant doit souligner que deux affirmations, opinions ou thèses ne se valent pas mais, au contraire, que l'une est meilleure que l'autre pour une raison donnée. Cette raison peut être centrée sur les faits (une thèse est supportée par des faits, l'autre non) ou par les méthodes ayant produit les preuves à l'appui de chaque thèse (par exemple, l'une s'appuie sur un faisceau d'observations, l'autre sur quelques témoignages anecdotiques), ou les sources qui les relaient. L'enseignant saisit donc, au sein de son enseignement classique, les opportunités pour apprendre aux élèves des stratégies pour mieux penser et mieux évaluer les informations (contenus et sources).

Cet enseignement est évolutif. Par exemple, évaluer l'expertise d'une source ne se fait pas de la même manière à sept et quinze ans. Dans un premier temps, l'enseignant cherche à éveiller les élèves à l'idée que des indices peuvent nous guider de manière plus ou moins efficace dans l'évaluation de l'expertise de la source. Plus tard, il peut évoquer le fonctionnement des sciences, les différents types d'expertise, etc.
Même avec tout cela, la difficulté du transfert ne doit pas être sous-estimée. S'il est trivial de considérer que plusieurs observations valent mieux qu'une, ou qu'une source experte vaut mieux qu'une source non experte, s'exercer à déceler la présence ou l'absence de tels critères dans des situations de la vie quotidienne est une véritable gageure. De telles attitudes restent coûteuses à mettre en place dans des contextes réels où nous perdons de vue leur intérêt et où nous cherchons souvent à aller vite. Seul un apprentissage régulier et explicite des mêmes considérations dans de très nombreuses situations permettrait d'espérer un transfert efficace dans des situations complexes de la vie quotidienne. Comment l'organiser ?

À l'école primaire comme au collège et au lycée, une solution est représentée par les connexions explicites entre disciplines. Par exemple, l'enseignant d'une discipline réalise son cours de manière classique, puis, à la fin du cours, il réserve un temps dédié pour des considérations de portée plus générale sur les preuves ou les sources à l'appui d'une thèse énoncée dans le cours. Il utilise un premier recours favorisant le transfert : l'explicitation. Il porte l'attention des élèves une première fois, par un discours de portée générale, sur la structure profonde du problème, au-delà du contexte dans lequel il a été rencontré. Un peu plus tard, l'enseignant d'une autre discipline (ou le même enseignant bien sûr dans le cadre de l'enseignement primaire) réalise à son tour un cours portant sur un contenu bien précis, différent du premier. À son tour, il consacre également quelques minutes pour présenter un message explicite et de portée plus générale. Il montre également le lien avec ce que les élèves ont découvert quelques jours plus tôt, favorisant le transfert vers un autre contexte. Les élèves reçoivent donc tour à tour un apprentissage ancré sur des cas concrets et un discours qui les aide à visualiser le problème à un niveau plus profond. Les enseignants peuvent aller plus loin en proposant aux élèves des situations (notamment sous forme d'évaluation formative comme des quiz) de la vie quotidienne qui mobilisent les mêmes obstacles et les

mêmes solutions. Les élèves peuvent aussi être incités à trouver eux-mêmes de tels exemples. Cette étape est critique : non seulement elle permet de viser le travail de la compétence dans le domaine visé à la fin (la vie quotidienne) mais elle aide également les élèves à comprendre ce que l'enseignant attend d'eux et à focaliser leur attention sur le message visé, au-delà des contextes d'apprentissage rencontrés.

Enfin, tout cet apprentissage ne peut se faire sans garde-fou. Il ne s'agit pas de faire croire aux élèves qu'ils doivent exercer aveuglément un esprit de doute et de critique sur chaque information. Une telle attitude deviendrait rapidement complètement contre-productive : en doutant aveuglément de tout, aussi bien des internautes et des blogueurs que de notre enseignant ou de notre médecin, nous n'adoptons pas un esprit critique efficace. L'enseignement de l'esprit critique comprend donc aussi une éducation quant à la manière de déployer les stratégies acquises à bon escient.

5.2 À faire

L'éducation de l'EC – tel que défini dans la deuxième partie de ce rapport – s'appuie sur trois principes (qui ont émergé dans la troisième partie) :

- outiller l'EC naturel des élèves grâce à des connaissances et critères appropriés au contexte et progressivement plus sophistiqués permettant d'évaluer la qualité épistémique des informations à leur disposition (plausibilité, pertinence, preuves à l'appui, qualité des sources) d'une manière plus avancée, voire experte ;
- amener les élèves à utiliser leur confiance métacognitive de façon explicite afin de bien calibrer la confiance que ces informations (ou les décisions qui se basent sur elles) méritent, sur la base de l'évaluation des preuves à l'appui, de la qualité des sources, de la pertinence et de la plausibilité des informations disponibles ;
- fournir aux élèves autant d'éléments que possible pour parvenir à cette évaluation et à ce calibrage. En plus des critères cités :
    - une plus vaste et profonde base de connaissances permettant de mieux évaluer la plausibilité des informations mais aussi de bien ajuster leur confiance et limiter les biais métacognitifs,
    - l'identification des circonstances plus à même de les mettre en difficulté et par rapport auxquelles ils doivent être particulièrement vigilants dans leurs évaluations et jugements. Ces circonstances motivent le recours à des critères et à une réflexion métacognitive explicite.

Les considérations concernant le transfert des apprentissages (présentées dans la quatrième partie) et d'autres d'ordre pratique permettent que l'approche pédagogique proposée ici amène les élèves à exploiter les contenus déjà enseignés et les compétences travaillées dans différentes disciplines, pour qu'ils comprennent, de manière explicite, que ces contenus disciplinaires utilisent des stratégies pouvant guider l'évaluation des informations. En effet, les programmes de l'Éducation nationale en France couvrent largement les besoins d'enseignement de l'esprit critique, en termes de compétences et de contenus enseignés. Ce qui manque, à notre avis, est :

- une approche assumée, explicite de l'enseignement de l'EC ;
- une liste des critères utiles pour améliorer l'évaluation de l'information ;
- des indications pratiques pour travailler ces critères et connaissances en relation avec l'EC, en identifiant les emplacements dans les curricula où ces critères et connaissances sont déjà cités et présents, sans donc altérer ces derniers (sans

introduire des heures d'enseignement dédiées à l'enseignement de l'esprit critique) mais en les tournant dans la direction d'une éducation de l'EC.

Les indications présentées ici ne se limitent cependant pas au contexte de l'éducation en France : les principes peuvent être adaptés à tout autre système éducatif, en raison des contenus et des compétences mis en avance par ce dernier.

Il s'agit ainsi d'aider les élèves à comprendre en quoi certains critères – introduits en cours de sciences, d'histoire, de littérature, de mathématique, de sport ou autre – sont pertinents pour évaluer la valeur épistémique d'une information (les preuves à l'appui de cette information ou la source qui porte l'information), la qualité d'une source ou la pertinence et plausibilité d'une information, et la manière de les mobiliser à bon escient. Puis de les inviter à se placer sur un continuum de confiance par rapport à l'information évaluée pour qu'ils prennent position de manière plus réflexive.

5.2.1 Outiller l'EC naturel. Comment choisir les critères à enseigner ?

Une liste détaillée des critères permettant d'évaluer n'importe quel type d'information dans n'importe quelle circonstance de façon experte serait probablement infinie. Dans chaque domaine de connaissance, pour chaque problématique, les experts du domaine cherchent à développer des critères toujours plus fins d'évaluation de l'information. Devons-nous alors renoncer à enseigner ces critères puisqu'il est impossible de rendre les élèves experts sur chaque question ? En réalité, la plupart de ces critères experts sont inadaptés pour des situations de la vie quotidienne et les contenus que nous y croisons. Les enseignements primaire et secondaire peuvent donc se focaliser sur des critères répondant à deux conditions pragmatiques :

- ils permettent d'évaluer de manière plus avancée et efficiente des contenus régulièrement rencontrés dans la vie quotidienne ou dans le contexte des connaissances fournies par l'école ;
- ils sont faciles à enseigner à partir des contenus disciplinaires et compétences présents dans les programmes scolaires.

Ainsi, par exemple, tout citoyen jeune et moins jeune est exposé à des affirmations basées sur des constats anecdotiques, non étayés, basés sur des observations fortuites. Quelle confiance donner à ces affirmations ? Cette question a un poids considérable dans la vie quotidienne. Tout citoyen est également exposé à et fabrique de façon presque automatique des explications causales à partir d'indices relativement généraux. Comment établir qu'une affirmation causale a des meilleures chances d'être correcte qu'une autre ? Sur la base de quels critères et à l'aide de quels outils pouvons-nous sortir de l'incertitude entre deux hypothèses alternatives ?

Une base de départ pour répondre à ces questions est constituée par ce que nous faisons naturellement, spontanément, quand nous sommes confrontés à une nouvelle information, à une affirmation :

- nous jugeons sa plausibilité, sa cohérence par rapport à l'ensemble de connaissances que nous possédons ;
- éventuellement, nous évaluons les éléments qui l'accompagnent : il peut s'agir d'arguments qui s'ajoutent à l'information, qui viennent la compléter ou la justifier et que nous évaluons d'un point de vue logique par rapport à leur pertinence en relation à l'information ;
- nous nous interrogeons sur la fiabilité de la source qui nous fournit l'information : son identité, sa bienveillance à notre égard, son absence d'intérêt à nous mentir ou

manipuler (caractère désintéressé de la source) ; ses compétences dans le domaine (expertise de la source) ;
- éventuellement nous nous posons la question de la manière avec laquelle les preuves à l'appui de l'affirmation ont été établis : s'agit-il d'une observation directe ? d'un souvenir ? etc.

L'objectif de l'éducation à l'esprit critique va être celui de s'ancrer sur ces critères naturels de jugement (voir la partie 3 de ce rapport) et d'améliorer leur utilisation en les rendant plus performants, plus adaptés pour bien juger la véracité d'une affirmation ou d'une information. On vise un niveau avancé d'utilisation de ces critères (celui spontané étant à considérer comme le niveau basique), et non un niveau expert, propre à certaines professions (journaliste, historien, scientifique de profession).

Pour cela, il s'agit de :

- transmettre des connaissances pour améliorer le jugement de **plausibilité** ;
- apprendre l'identification d'un **bon argument**, mais surtout des éléments qui peuvent nous convaincre à tort (leviers de persuasions, éléments saillants et disponibles, envie de croire, envie de lever l'incertitude...)
- apprendre l'identification des **sources** qui méritent notre confiance, dans différents contextes, en s'appuyant sur l'absence d'intérêts privés, et la compétence de la source;
- apprendre la reconnaissance de **preuves** de meilleure qualité sur la base des méthodes ayant permis de les obtenir (épistémologie du premier ordre).

5.2.2 Où trouver ces critères dans les programmes de l'éducation française ?

Où trouver précisément de tels critères dans les enseignements disciplinaires ? Les enseignements scientifiques, par exemple, proposent régulièrement de tester une hypothèse (une idée reçue fausse, une intuition basée sur des observations, etc.). L'enseignant peut présenter la tâche sous la forme d'une évaluation du bien-fondé de l'idée. Il met alors en place une procédure pour juger si l'information en question est très fiable, plausible ou probablement fausse. Il peut donc choisir de présenter le même cours en insistant sur la tâche qui incombe aux élèves, c'est-à-dire chercher des preuves en faveur de l'hypothèse (ou de son opposé). La procédure scientifique est ainsi présentée comme un ensemble d'outils permettant de produire des preuves (et donc des connaissances fiables).

Prenons un exemple concret : nous sommes à Vienne, au XIXe siècle. Dans un hôpital se déclenche une épidémie touchant les femmes enceintes : sont accusés tour à tour le passage du prêtre, les pratiques médicales et le manque de fenêtres. Comment trancher ? L'activité pédagogique proposée permet aux élèves d'accumuler les preuves en faveur de chaque hypothèse et d'acquérir des critères afin d'évaluer la fiabilité des preuves récoltées. À la fin, les élèves doivent conseiller au directeur de l'hôpital une décision fondée sur les meilleures preuves disponibles.

Reconnaissons que, dans la vie quotidienne, il nous est souvent impossible de mobiliser de tels outils. Cependant, nous pouvons les utiliser comme des critères pour distinguer une information de qualité d'une information qui ne mérite pas notre confiance. Prenons un autre exemple : un protocole scientifique s'appuie sur des observations répétées car le hasard peut produire des événements particuliers que nous interpréterons alors à tort. Si nous testons l'effet d'un médicament sur un seul patient, celui-ci peut guérir sans lien avec le médicament, mais pour bien d'autres raisons. Deux ou trois patients ne suffisent pas non plus à produire une connaissance fiable sur le médicament : il est nécessaire e multiplier les observations. Cet outil de la science peut devenir un critère d'évaluation de l'information pour le citoyen. Ainsi,

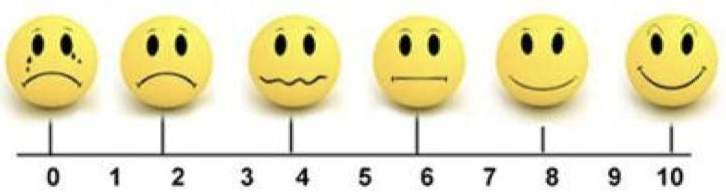
un citoyen confronté à des affirmations contradictoires – l'une appuyée sur un témoignage personnel unique et l'autre soutenue par de multiples observations – devrait accorder sa confiance à la seconde affirmation sur la base du critère « observations multiples ».

De tels critères sont faciles à trouver dans les enseignements scientifiques au sens restreint du terme (sciences de la vie et de la Terre, sciences physiques et chimiques) car la méthode scientifique est une accumulation d'outils permettant de produire la meilleure connaissance possible. L'enseignant peut alors souligner qu'il est important de multiplier les observations, de ne pas confondre cause et corrélation, de ne faire varier qu'un facteur à la fois pour chercher une cause, etc. Les temps d'analyse de documents, de démarche d'investigation ou de débat sont favorables pour illustrer les notions de faits, de preuves et de fiabilité. Soulignons également l'importance de faire comprendre aux élèves la nature de l'entreprise scientifique pour elle-même, en tant que mécanisme social de production de la connaissance. Le fonctionnement des institutions formées d'experts qui produisent et évaluent la connaissance est un savoir utile à tout citoyen pour se forger une confiance éclairée envers le savoir scientifique.

Cependant, il est possible d'aborder l'importance d'évaluer l'information sur la base de critères dans toutes les disciplines :

- en mathématiques, avec l'enseignement des statistiques et des probabilités : l'enseignant peut profiter de toutes les situations où les élèves étudient des jeux de données et essaient d'en extraire du sens pour travailler sur leur perception du hasard, l'importance de la multiplication des observations, le recours à des outils de traitement plus objectifs (comme la moyenne) tout en restant vigilants sur l'interprétation des données en l'absence de connaissances, etc. ;
- en technologie, où la démarche technologique s'associe de près à celle scientifique : l'enseignant peut saisir toutes les occasions où sont émises des hypothèses à l'épreuve des faits pour évoquer la différence entre de simples opinions et des connaissances fiables sur lesquelles appuyer ses décisions ;
- en histoire et en géographie, autres disciplines scientifiques où les élèves analysent régulièrement les preuves et les sources qui étayent des thèses ;
- en éducation physique et sportive, où nous pouvons vouloir vérifier des affirmations notamment sur nos compétences sportives et la manière de les améliorer ;
- dans les enseignements littéraires : il est possible de chercher dans différents textes ce qui relève du fait, de la description, de la preuve, de l'argumentation et de l'interprétation. Ces textes ont des objectifs différents et il est indispensable de comprendre les intentions de l'auteur. L'enseignant peut également confronter le point de vue de cultures différentes sur un même phénomène, voire sur les autres cultures, pour discuter de la subjectivité de certaines opinions, de l'origine de certains stéréotypes, etc. ;
- dans les enseignements artistiques, les élèves peuvent questionner le rôle des sens dans la perception. Comme pour l'étude des textes, ils peuvent chercher dans différents œuvres et supports l'intention de l'artiste de retranscrire la réalité (comme dans un dessin naturaliste, anatomique) ou, au contraire, de traduire une perception subjective.

5.2.3 Exercer la confiance métacognitive de façon explicite. Se placer sur un continuum de confiance et réfléchir à sa signification



Face à une information, nous pouvons nous sentir plus ou moins confiants que son contenu est correct. Ce sentiment peut être explicité, et nous pouvons réfléchir sur ce qui nous amène à considérer l'information comme fiable et digne de notre confiance ou, à l'inverse, douteuse. Les critères que nous avons présentés plus haut ont cet objectif. Mais il faut les rendre opérationnels, c'est-à-dire faire en sorte qu'ils influencent notre confiance en une information, affirmation et décision. C'est pour cette raison que nous introduisons ici le concept, et l'outil, du continuum de confiance. Les élèves réfléchissent à la qualité des preuves et à la fiabilité des sources à l'appui d'une position. Même si les preuves sont faibles et les sources ne sont pas d'exceptionnelle qualité, cela ne les empêche pas de prendre une position. Mais ils peuvent réfléchir à l'incertitude qui accompagne cette prise de position. Comme nous l'avons discuté dans la troisième partie de ce rapport, l'incertitude en effet est une condition de changement d'opinion, et reconnaître l'incertitude qui accompagne nos positions peut faciliter la compréhension et éventuellement l'adoption de celles des autres. Le continuum de confiance permet donc plus facilement de modifier nos positions si nécessaire. Il permet également d'exprimer notre certitude et confiance, et donc de montrer que certaines opinions ne changeront pas si vite. Par exemple, le fait de se placer du côté droit du continuum face à l'affirmation : « La Terre a une forme plutôt sphérique » est une attitude qui relève de l'EC, à partir du moment où les sources qui l'affirment méritent notre confiance et/ou que les arguments et les preuves sont à notre portée.

5.2.4 Éléments additionnels pour parvenir à une juste évaluation de la qualité de l'information et un calibrage de la confiance

5.2.4.1 Acquérir une plus vaste et profonde base de connaissances

Nous avons à plusieurs reprises souligné (dans les deuxième et quatrième parties de ce rapport notamment) le fait que l'EC n'est pas disjoint des connaissances et des contenus sur lesquels il s'exerce. Si, d'un côté, l'ancrage dans les contenus limite le transfert des outils acquis à de nouveaux contextes, de l'autre, une plus vaste et profonde base de connaissances permettrait de mieux évaluer la plausibilité des informations disponibles. Ainsi, de nombreuses personnes seraient probablement prêtes à réagir face à la révélation selon laquelle dans nos robinets coule un liquide dangereux, dont l'inhalation directe peut causer la mort, liquide qui d'ailleurs est utilisé couramment dans les systèmes de refroidissement des centrales nucléaires, et qui représente de plus la composante principale des pluies acides : il s'agit du monoxyde de dihydrogène. Les réactions seraient vite estompées par la connaissance suivante : le monoxyde de dihydrogène (H_20) est en réalité l'eau. Une connaissance factuelle permet ici d'évaluer la portée d'une information et notamment la plausibilité des affirmations avancées concernant ses risques. De manière plus générale, posséder des connaissances concernant l'énergie et les formes d'énergie connues permet de mieux jauger la plausibilité d'affirmations sur les énergies psychiques ou d'autres affirmations contredisant de manière évidente une masse importante de connaissances actuelles en physique ou en biologie. Posséder des connaissances factuelles augmente également les chances de savoir identifier les sources d'informations les plus informées et fiables dans un domaine, et cela permet de mieux calibrer notre confiance et notamment de contrecarrer l'effet de biais métacognitif qui se manifeste plus particulièrement par rapport à nos zones d'ignorance.

L'enseignement qui enseigne aux élèves une image vaste et connectée des principales lois de la nature et des phénomènes physiques et biologiques connus leur fournit également une base vaste et solide pour qu'ils puissent jauger de la plausibilité d'une quantité d'informations. Cela ne signifie pas que d'autres informations pour lesquelles les connaissances développées par l'individu sont faibles ne seront pas jugées de manière plus naïve. La dépendance de l'EC des contenus de connaissance nous rend inévitablement faillibles. L'objectif est d'améliorer les capacités de l'EC en l'outillant, pas d'atteindre une idéale perfection.

5.2.4.2 Prendre conscience des situations où l'exercice de l'EC avancé est requis et des circonstances plus à même de nous mettre en difficulté

Le déploiement des critères avancés d'évaluation de la qualité de l'information et l'exercice volontaire et explicite des jugements métacognitifs de confiance exigent un effort et comportent un coût par rapport au recours de critères plus intuitifs qui ne passent pas par une acquisition de connaissances ni par des évaluations implicites.

Il ne faudrait pas donner l'impression que la mobilisation de ces critères est une nécessité constante, et que tout moment de la vie va se transformer en un lourd exercice de réflexion. Ces conditions ne sont mobilisées que quand la situation le nécessite : un contenu nouveau, une décision importante à prendre, une opinion qui, une fois fixée, peut comporter des conséquences non anodines pour soi ou pour les autres.

Motiver le recours à de tels critères est cependant également un défi. En effet, nous savons que des observations multiples sont plus fiables que des observations limitées, ou que des sources cachées sont suspectes. Mais nous avons également à notre disposition des stratégies d'évaluation intuitives, souvent moins coûteuses, surtout dans des contextes « écologiques » chargés comme ceux que nous rencontrons dans notre vie quotidienne. Nous devons parfois prendre des décisions de manière rapide, dans des situations où d'autres enjeux sont impliqués (le regard des autres, des émotions, etc.). Nous ne devons donc pas sous-estimer la difficulté de l'effort que nous devons poursuivre pour rendre accessible le recours à ces critères. Un levier sur lequel jouer concerne donc l'explicitation des limites de nos outils naturels. L'enseignant peut créer des situations pédagogiques qui amènent les élèves à répondre de manière intuitive à une problématique pour déceler les défauts de telles stratégies. Par exemple, si plusieurs groupes d'élèves parviennent, après un échantillonnage limité, à des conclusions différentes, l'enseignant peut rebondir en leur montrant l'importance de ne pas céder à la tentation naturelle de conclure à partir d'observations limitées. La stratégie alternative, plus coûteuse (s'appuyer sur des observations multiples), est ainsi renforcée.

L'intérêt de s'appuyer sur des situations concrètes est clair ici : plutôt que de transmettre un discours abstrait et des conseils généraux qui pourraient paraître triviaux (« des observations répétées valent mieux qu'une observation isolée »), la situation pédagogique favorise une véritable réflexion de l'élève sur ses propres stratégies et sur les méthodes qui produisent des conclusions fiables ou non. Ces situations ancrées sur des contextes réalistes pour l'enfant ont plus de chance de lui permettre de développer des compétences d'évaluation de l'information, si elles sont par la suite associées à un discours cette fois plus abstrait (« des conclusions issues d'observations répétées sont plus fiables en principe que celles basées sur une ou deux observations »). Cette formulation abstraite a plus de chances d'être mémorisée et permettrait d'être mobilisée par la suite dans d'autres contextes. Cependant, elle n'aurait du sens que si elle avait été associée auparavant à un apprentissage dans un cadre pertinent au regard de l'élève.

5.2.5 Tester

Dans cette partie, nous avons cherché à fournir des principes généraux d'éducation de l'EC qui découlent, selon nous, des analyses et revues de littérature présentées dans ce rapport. Aucun principe, quoique fondé sur des connaissances étayées, ne peut garantir l'efficacité d'une approche pédagogique dans des conditions écologiques, de classe. Les connaissances que nous avons mobilisées nous semblent néanmoins justifier l'effort et l'opportunité de tester l'efficacité d'interventions éducatives qui s'en inspirent. Des données empiriques, de terrain, pourront nous en dire plus sur la manière d'envisager le plus efficacement possible l'éducation de l'EC.

Annexes

Notes

[i] *"Critical thinking is a widely accepted educational goal. Its definition is contested, but the competing definitions can be understood as differing conceptions of the same basic concept: careful thinking directed to a goal. Conceptions differ with respect to the scope of such thinking, the type of goal, the criteria and norms for thinking carefully, and the thinking components on which they focus. Its adoption as an educational goal has been recommended on the basis of respect for students' autonomy and preparing students for success in life and for democratic citizenship. 'Critical thinkers' have the dispositions and abilities that lead them to think critically when appropriate. The abilities can be identified directly; the dispositions indirectly, by considering what factors contribute to or impede exercise of the abilities. Standardized tests have been developed to assess the degree to which a person possesses such dispositions and abilities. Educational intervention has been shown experimentally to improve them, particularly when it includes dialogue, anchored instruction, and mentoring. Controversies have arisen over the generalizability of critical thinking across domains, over alleged bias in critical thinking theories and instruction, and over the relationship of critical thinking to other types of thinking."*

Hitchcock 2018

[ii] *"It is daunting for someone who is becoming acquainted with the field of critical thinking to find the assortment of definitions provided by long-time participants in the field. It might seem from this assortment of definitions that the field of critical thinking is in chaos. But I do not think it is. What seems more reasonable to me is that the set of definitions in Table 1 below are different descriptions of the same concept… which I hold is the mainstream concept of critical thinking. ...*

These are definitions with which over the years I have become familiar while pursuing my interests in critical thinking."
*1) "**Active, persistent, and careful consideration** of any belief or supposed form of knowledge in the light of the **grounds that support it** and the further conclusions to which it tends" (Dewey, 1933, p. 9 (first edition 1910).*
*2) "The ability to think critically ...involves three things: (1) an attitude of being disposed to **consider in a thoughtful way** the problems and subjects that come within the range of one's experiences, (2) knowledge of the methods of logical inquiry and reasoning, and (3) some skill in applying those methods (Glaser 1941).*
*3) "Critical thinking is the intellectually disciplined process of **actively and skillfully conceptualizing, applying, analyzing, synthesizing,** and/or **evaluating information** gathered from, or generated by, observation, experience, reflection, reasoning, or communication, as a guide to belief and action" (Scriven and Paul 1987).*
*4) "Critical thinking is reasonable **reflective thinking** that is focused on **deciding what to believe or** do" (Ennis 1987a; 1987b; 1991; 2011, 2015).*
*5) "A critical thinker is one who is appropriately moved by **reasons**" (Siegel 1988, p. 32).*
*6) "Skillful, responsible thinking that is conducive to good judgment because it is sensitive to context, **relies on criteria, and is self-correcting**" (Lipman 1988).*
*7) "**Purposeful**, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual **considerations upon which that judgment is based"** (Facione 1990, Table 1).*
*8) "Critical thinking is skilled, **active** interpretation and **evaluation of observations, communications, information, and argumentatio**n as a guide to thought and action" (Fisher and Scriven 1997, p. 20)1.*
*9) "The practice of **identifying, having, and giving good reasons** for one's beliefs, values, and actions, given one's goals of truth and avoidance of error" (Possin, 2002).*
*10) "Thinking that attempts to arrive at a judgment only after **honestly evaluating** alternatives with respect to available evidence and arguments" (Hatcher and Spencer, 2006, p. 1).*
*11) "The **careful examination** of an issue in order to reach a reasoned judgment" (Bailin and Battersby, 2010).*
*12) "The **careful, deliberate determination of whether we should accept, reject, or suspend judgment** about a claim, and the degree of confidence with which we accept or reject it" (Moore and Parker, 2014).*

*13) "The **articulated judgment** of an intellectual product arrived at on the basis of plus-minus considerations of the product in terms of appropriate standards (or criteria)" (Johnson, 2014).*
*14) "[**Reasonable**] inquiry and argument" (Kuhn, 2015, p. 47)*

"Non-academic internet-dictionary definitions of 'critical thinking', which I assume have captured current use of the term in the media, also express the concept of critical thinking."

1) "Disciplined thinking that is clear, rational, open-minded, and informed by evidence" (Dictionary.com, 2016)
2) "The objective analysis and evaluation of an issue in order to form a judgment"(Oxforddictionaries.com, 2016)
3) "A mental process of reviewing clear, rational thoughts based on evidence to reach an answer or a conclusion (Yourdictionary.com, 2016)."

Ennis 2016

[3] "*Although many psychologists and others have proposed definitions for the term 'critical thinking', these definitions tend to be similar with considerable overlap among definitions. In a review of the critical thinking literature, Fischer & Spiker (2000) found that most definitions for the term 'critical thinking' include reasoning/logic, judgment, metacognition, reflection, questioning and mental processes. Jones and his colleagues (Jones, et al. 1995a, b) obtained consensus from among 500 policy makers, employers, and educators who agree that critical thinking is a broad term that describes reasoning in an open-ended manner and with an unlimited number of solutions. It involves constructing a situation and supporting the reasoning that went into a conclusion."*

Halpern 2013

[4] "*Judging by the attention given to critical thinking in educational journals and in the social documents of governing agencies, support for teaching critical thinking at all levels of education is extremely strong in North America and the UK. But, agreement about teaching critical thinking persists only so long as theorists remain at the level of abstract discussion and permit their use of the term to remain vague. As soon as they begin to spell out in more concrete terms what critical thinking consists in, what educational attainments are required if one is to be a critical thinker, and what means are likely to be efficacious in teaching persons to think critically, that is to say, as soon as they interpret the term in such a way as to provide a clear conception of critical thinking, agreement evaporates*."

Bailin et al. 1999

[5] "*The experts articulated an ideal. It may be that no person is fully adept at the skills and sub-skills the experts found to be central to CT. It may be that no person has fully cultivated all the affective dispositions which characterize a good critical thinker. It may be that humans compartmentalize their lives in ways that CT is more active and evident in some areas than in others. This gives no more reason to abandon the effort to infuse CT into the educational system than that knowing no friendship is perfect gives reason to despair of having friends. The experts' proposal in putting the ideal before education community is that it should serve as a rich and worthy goal guiding CT assessment and curriculum development at all educational levels. We understand critical thinking to be the purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based. CT is essential as a tool of inquiry. As such, CT is a liberating force in education and a powerful resource in one's personal and civic life. While not synonymous with good thinking, CT is a pervasive and self-rectifying human phenomenon. The ideal critical thinker is habitually inquisitive, well-informed, trustful of reason, open-minded, flexible, fair-minded in evaluation, honest in facing personal biases, prudent in making judgments, willing to reconsider, clear about issues, orderly in complex matters, diligent in seeking relevant information, reasonable in the selection of criteria, focused in inquiry, and persistent in seeking results which are as precise as the subject and the circumstances of inquiry permit. Thus, educating good critical thinkers means working toward this ideal. It combines developing CT skills with nurturing those dispositions which consistently yield useful insights and which are the basis of a rational and democratic society."*

Facione 1990

[6] "*Individuals vary in their views of what students should be taught … But there is no disagreement on the importance of critical thinking skills. In free societies, the ability to think critically is viewed as a cornerstone of*

individual civic engagement and economic success. We may disagree about which content students should learn, but we at least agree that, whatever they end up learning, students ought to think critically about it. … Despite this consensus it is not clear we know what we mean by 'critical thinking'. I will offer a commonsensical view (Willingham, 2007). You are thinking critically if (1) your thinking is novel - that is, you aren't simply drawing a conclusion from a memory of a previous situation and (2) your thinking is self-directed - that is, you are not merely executing instructions given by someone else and (3) your thinking is effective - that is, you respect certain conventions that make thinking more likely to yield useful conclusions. These would be conventions like 'consider both sides of an issue', and 'offer evidence for claims made', and 'don't let emotion interfere with reason'. This last characteristic will be our main concern, and as we'll see, what constitutes effective thinking varies from domain to domain."

Willingham 2019

[7] *"The intellectual roots of critical thinking are as ancient as its etymology, traceable, ultimately, to the teaching practice and vision of Socrates 2,500 years ago who discovered by a method of probing questioning that people could not rationally justify their confident claims to knowledge. Confused meanings, inadequate evidence, or self-contradictory beliefs often lurked beneath smooth but largely empty rhetoric. Socrates established the fact that one cannot depend upon those in 'authority' to have sound knowledge and insight. He demonstrated that persons may have power and high position and yet be deeply confused and irrational. He established the importance of asking deep questions that probe profoundly into thinking before we accept ideas as worthy of belief. He established the importance of seeking evidence, closely examining reasoning and assumptions, analyzing basic concepts, and tracing out implications not only of what is said but of what is done as well. His method of questioning is now known as 'Socratic Questioning' and is the best known critical thinking teaching strategy. In his mode of questioning, Socrates highlighted the need in thinking for clarity and logical consistency. …*

Socrates' practice was followed by the critical thinking of Plato (who recorded Socrates' thought), Aristotle, and the Greek skeptics, all of whom emphasized that things are often very different from what they appear to be and that only the trained mind is prepared to see through the way things look to us on the surface (delusive appearances) to the way they really are beneath the surface (the deeper realities of life). From this ancient Greek tradition emerged the need, for anyone who aspired to understand the deeper realities, to think systematically, to trace implications broadly and deeply, for only thinking that is comprehensive, well-reasoned, and responsive to objections can take us beyond the surface. …

*In the Middle Ages, the tradition of systematic critical thinking was embodied in the writings and teachings of such thinkers as Thomas Aquinas (**Sumna Theologica**) who to ensure his thinking met the test of critical thought, always systematically stated, considered, and answered all criticisms of his ideas as a necessary stage in developing them. … In the Renaissance (15th and 16th Centuries), a flood of scholars in Europe began to think critically about religion, art, society, human nature, law, and freedom. … Francis Bacon, in England, was explicitly concerned with the way we misuse our minds in seeking knowledge. He recognized explicitly that the mind cannot safely be left to its natural tendencies. In his book **The Advancement of Learning**, he argued for the importance of studying the world empirically. He laid the foundation for modern science with his emphasis on the information-gathering processes. He also called attention to the fact that most people, if left to their own devices, develop bad habits of thought (which he called 'idols') that lead them to believe what is false or misleading. ...*

*Some fifty years later in France, Descartes wrote what might be called the second text in critical thinking, **Rules For the Direction of the Mind**. In it, Descartes argued for the need for a special systematic disciplining of the mind to guide it in thinking. He articulated and defended the need in thinking for clarity and precision. He developed a method of critical thought based on the principle of systematic doubt. …*

*Hobbes and Locke (in 16th and 17th Century England) displayed the same confidence in the critical mind of the thinker that we find in Machiavelli. Neither accepted the traditional picture of things dominant in the thinking of their day. Neither accepted as necessarily rational that which was considered 'normal' in their culture. Both looked to the critical mind to open up new vistas of learning. Hobbes adopted a naturalistic view of the world in which everything was to be explained by evidence and reasoning. Locke defended a common-sense analysis of everyday life and thought. He laid the theoretical foundation for critical thinking about basic human rights and the responsibilities of all governments to submit to the reasoned criticism of thoughtful citizens. It was in this spirit of intellectual freedom and critical thought that people such as Robert Boyle (in the 17th Century) and Sir Isaac Newton (in the 17th and 18th Century) did their work. In his **Sceptical Chymist**, Boyle severely criticized*

the chemical theory that had preceded him. Newton, in turn, developed a far-reaching framework of thought which roundly criticized the traditionally accepted world view. He extended the critical thought of such minds as Copernicus, Galileo, and Kepler. After Boyle and Newton, it was recognized by those who reflected seriously on the natural world that egocentric views of world must be abandoned in favor of views based entirely on carefully gathered evidence and sound reasoning. …

Another significant contribution to critical thinking was made by the thinkers of the French Enlightenment: Bayle, Montesquieu, Voltaire, and Diderot. They all began with the premise that the human mind, when disciplined by reason, is better able to figure out the nature of the social and political world. What is more, for these thinkers, reason must turn inward upon itself, in order to determine weaknesses and strengths of thought. They valued disciplined intellectual exchange, in which all views had to be submitted to serious analysis and critique. They believed that all authority must submit in one way or another to the scrutiny of reasonable critical questioning."

Foundation for critical thinking, https://www.criticalthinking.org/pages/defining-critical-thinking/766

[8] *"The essence of critical thinking is suspended judgment; and the essence of this suspense is inquiry to determine the nature of the problem before proceeding to attempts at its solution. This, more than any other thing, transforms mere inference into tested inference, suggested conclusions into proof. ...*

Active, persistent and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it, and the further conclusions to which it tends, constitutes reflective thought."

Dewey 1933

[9] *"The real problem of intellectual education is the transformation of more or less casual curiosity and sporadic suggestion into attitudes of alert, cautious, and thorough inquiry."*

Dewey 1933

"Thinking is stoppage of the immediate manifestation of impulse until that impulse has been brought into connection with other possible tendencies to action so that a more comprehensive and coherent plan of activity is formed. Some of the other tendencies to action lead to use of eye, ear, and hand to observe objective conditions; others result in recall of what has happened in the past. Thinking is thus a postponement of immediate action, while it effects internal control of impulse through a union of observation and memory, this union being the heart of reflection. What has been said explains the meaning of the well-worn phrase 'self-control'. The ideal aim of education is creation of power of self-control."

Dewey 1997

[10] *1. Transit: "The other day, when I was downtown on 16th Street, a clock caught my eye. I saw that the hands pointed to 12:20. This suggested that I had an engagement at 124th Street, at one o'clock. I reasoned that as it had taken me an hour to come down on a surface car, I should probably be twenty minutes late if I returned the same way. I might save twenty minutes by a subway express. But was there a station near? If not, I might lose more than twenty minutes in looking for one. Then I thought of the elevated, and I saw there was such a line within two blocks. But where was the station? If it were several blocks above or below the street I was on, I should lose time instead of gaining it. My mind went back to the subway express as quicker than the elevated; furthermore, I remembered that it went nearer than the elevated to the part of 124th Street I wished to reach, so that time would be saved at the end of the journey. I concluded in favor of the subway, and reached my destination by one o'clock."*

2. Ferryboat: "Projecting nearly horizontally from the upper deck of the ferryboat on which I daily cross the river is a long white pole, bearing a gilded ball at its tip. It suggested a flagpole when I first saw it; its color, shape, and gilded ball agreed with this idea, and these reasons seemed to justify me in this belief. But soon difficulties presented themselves. The pole was nearly horizontal, an unusual position for a flagpole; in the next place, there was no pulley, ring, or cord by which to attach a flag; finally, there were elsewhere two vertical staffs from which flags were occasionally flown. It seemed probable that the pole was not there for flag-flying. I then tried to imagine all possible purposes of such a pole, and to consider for which of these it was best suited: (a) Possibly it was an ornament. But as all the ferryboats and even the tugboats carried poles, this hypothesis was rejected. (6) Possibly it was the terminal of a wireless telegraph. But the same considerations made this improbable. Besides, the more natural place for such a terminal would be the highest part of the boat, on top of

the pilot house, (c) Its purpose might be to point out the direction in which the boat is moving. In support of this conclusion, I discovered that the pole was lower than the pilot house, so that the steers man could easily see it. Moreover, the tip was enough higher than the base, so that, from the pilot's position, it must appear to project far out in front of the boat. Moreover, the pilot being near the front of the boat, he would need some such guide as to its direction. Tug boats would also need poles for such a purpose. This hypothesis was so much more probable than the others that I accepted it. I formed the conclusion that the pole was set up for the purpose of showing the pilot the direction in which the boat pointed, to enable him to steer correctly."

3. Bubbles: "In washing tumblers in hot soap suds and placing them mouth downward on a plate, I noticed that bubbles appeared on the outside of the mouth of the tumblers and then went inside. Why? The presence of bubbles suggests air, which I note must come from inside the tumbler. I see that the soapy water on the plate prevents escape of the air save as it may be caught in bubbles. But why should air leave the tumbler? There was no substance entering to force it out. It must have expanded. It expands by increase of heat or by increase of pressure, or by both. Could the air have become heated after the tumbler was taken from the hot suds? Clearly not the air that was already entangled in the water. If heated air was the cause, cold air must have entered in transferring the tumblers from the suds to the plate. I test to see whether this supposition is true by taking several more tumblers out. Some I shake so as to make sure of entrapping cold air in them. Some I take out, holding them mouth downward in order to prevent cold air from entering. Bubbles appear on the outside of every one of the former and on none of the latter. I must be right in my inference. Air from the outside must have been expanded by the heat of the tumbler, which explains the appearance of the bubbles on the outside. But why do they then go inside? Cold contracts. The tumbler cooled and also the air inside it. Tension was removed, and hence bubbles appeared inside. To be sure of this, I test by placing a cap of ice on the tumbler while the bubbles are still forming outside. They soon reverse."

Dewey 1933

[11] *"1. An attitude of being disposed to consider in a thoughtful way the problems and subjects to come within the range of one's own experiences; 2. Knowledge of the methods of logical inquiry and reasoning; 3. Some skills in applying those methods. Critical thinking calls for a persistent effort to examine any belief or supposed form of knowledge in the light of the evidence that supports it and the further conclusions to which it tends. It also generally requires ability to recognise problems, to find workable means for meeting this problems, to gather and marshal pertinent information, to recognise unstated assumptions and values, to comprehend and use language with accuracy, clarity and discrimination, to interpret data, to appraise evidence and evaluate arguments, to recognise the existence (or non-existence) of logical relationships between propositions, to draw warranted conclusions and generalizations at which one arrives, to reconstruct one's patterns of belief on the basis of wider experience, and to render accurate judgments about specific things and qualities in everyday life."*

Glaser 1941

[12] *"The application of critical thought to pedagogy in our schools was given a major impetus in the middle of this century. According to Cuban (1984), 'The work of B.O. Smith in the 1950s and subsequently Robert Ennis, have provided a scholarly rationale and specific ingredients for designing school programs to develop critical thought'."*

Lewis & Smith 1993

[13] *"The Philosophy for Children Program, developed by Matthew Lipman, represents one way to introduce critical thinking skills. … In Lipman's program, fifth and sixth grade children read and discuss a set of novels. Through these activities, students are encouraged to develop philosophical reasoning skills including commitments to impartiality and objectivity, relevance, consistency, and the search for defensible reasons for behavior. While Lipman's program is designed especially to be added to the curriculum, other philosophers incorporate critical thinking into the existing curriculum. For example, Paul, Binker, & Weil's Critical Thinking Handbook (1990) helps K-3 teachers remodel their lesson plans in language arts, social studies, and science in order to incorporate critical thinking."*

Lewis & Smith 1993

[14] *"It is true of teachers and farmers and theoretical physicists as well: all must make judgments in the practice of their occupations and in the conduct of their lives. There are practical, productive, and theoretical judgments, as Aristotle would have put it. Insofar as we make such judgments well, we can be said to behave wisely.(...)*

They are likely to be good judgments if they are the product of skillfully performed acts guided or facilitated by appropriate instruments and procedures."

Lipman 1987

"Thus architects will judge a building by using criteria such as utility, safety, and beauty. And presumably critical thinkers will rely upon such time-tested criteria such as validity, evidential warrant, and consistency. Any area of practice - architectural, cognitive and the like - should be able to cite the criteria by which the practice is guided."

Lipman 1987

[15] "*Critical thinking is a cultivation of that strand of traditional education which stresses the cultivation of wisdom and its application to both practice and life. To strengthen critical thinking in schools and colleges, it is necessary to know its defining features, its characteristic outcomes, and its underlying conditions. The outcomes of critical thinking are judgments; and the nature of judgment is such that critical thinking may be defined as skillful, responsible thinking that facilitates good judgment because (1) it relies upon criteria; (2) it is self-correcting, and (3) it is sensitive to context. The very meaning of 'criterion' is 'a rule or principle utilized in the making of judgments.' Judgment, in turn, is a skill; therefore critical thinking is skillful thinking, and skills can only be defined through criteria by which performance can be evaluated. So critical thinking is thinking that both employs criteria and can be assessed by appeal to criteria. Important criteria are reliability, strength, relevance, coherence, and consistency. Critical thinking is self-corrective, promoting a community of inquiry in the classroom by requiring students to discover weaknesses in their own thinking and to rectify faults in their procedures. Finally, thinking that is sensitive to context involves recognition of: exceptional or irregular circumstances and conditions; special limitations, contingencies, or constraints; overall configurations; the possibility that evidence is atypical; and the possibility that some meanings do not translate from one context or domain to another. Exemplary instances of critical thinking can be found in the best practice of law and medicine. The relevance of critical thinking to the enhancement of K-12 and college education is related to the shift from learning to thinking as the focus of education and to the goal of helping students develop the reasoning skills that will enable them to exercise good judgment.'*

Lipman, 1988

[16] " *'Critical thinking', according to Paul et al., 'is disciplined, self-directed thinking which exemplifies the perfections of thinking appropriate to a particular mode or domain of thought' (p. 361). Paul identifies the following perfections of thought: clarity, precision, specificity, accuracy, relevance, consistency, logic, depth, completeness, significance, fairness, and adequacy (p. 361)."*

Lewis & Smith 1993

[17] "*The eighties witnessed a growing accord that the heart of education lies exactly where traditional advocates of a liberal education always said it was — in the processes of inquiry, learning and thinking rather than in the accumulation of disjointed skills and senescent information. By the decade's end the movement to infuse the K-12 and post-secondary curricula with critical thinking (CT) had gained remarkable momentum. This success also raised vexing questions: What exactly are those skills and dispositions which characterize CT? What are some effective ways to teach CT? And how can CT, particularly if it becomes a campus-wide, district-wide or statewide requirement, be assessed? When asked by the individual professor or teacher seeking to introduce CT into her own classroom, such questions are difficult enough. But they take on social, fiscal, and political dimensions when asked by campus curriculum committees, school district offices, boards of education, and the educational testing and publishing industries. Given the central role played by philosophers in articulating the value, both individual and social, of CT, in analyzing the concept of CT, in designing college level academic programs in CT, and in assisting with efforts to introduce CT into the K-12 curriculum, it is little wonder that the American Philosophical Association, through its Committee on Pre-College Philosophy, took great interest in the CT movement and its impact on the profession. In December of 1987 that committee asked this investigator to make a systematic inquiry into the current state of CT and CT assessment."*

Facione 1990

[18] "*At one level we all know what 'critical thinking' means — it means good thinking, almost the opposite of illogical, irrational, thinking. But when we test our understanding further, we run into questions. For example, is critical thinking the same as creative thinking, are they different, or is one part of the other? How do critical*

thinking and native intelligence or scholastic aptitude relate? Does critical thinking focus on the subject matter or content that you know or on the process you use when you reason about that content? ...

Suggestion: What can the strong critical thinkers do (what mental abilities do they have), that the weak critical thinkers have trouble doing? What skills or approaches do the strong critical thinkers habitually seem to exhibit which the weak critical thinkers seem not to possess?

...look for a list of mental skills and habits of mind, the experts, when faced with the same problem you are working on, refer to their lists as including cognitive skills and dispositions. As to the cognitive skills here is what the experts include as being at the very core of critical thinking: interpretation, analysis, evaluation, inference, explanation, and self-regulation. …

Quoting from the consensus statement of the national panel of experts: ***interpretation*** *is 'to comprehend and express the meaning or significance of a wide variety of experiences, situations, data, events, judgments, conventions, beliefs, rules, procedures, or criteria.' Interpretation includes the sub-skills of categorization, decoding significance, and clarifying meaning. … analysis is 'to identify the intended and actual inferential relationships among statements, questions, concepts, descriptions, or other forms of representation intended to express belief, judgment, experiences, reasons, information, or opinions.' The experts include examining ideas, detecting arguments, and analyzing arguments as sub-skills of analysis. …*

evaluation *as meaning 'to assess the credibility of statements or other representations which are accounts or descriptions of a person's perception, experience, situation, judgment, belief, or opinion; and to assess the logical strength of the actual or intended inferential relationships among statements, descriptions, questions or other forms of representation.' … To the experts* ***inference*** *means 'to identify and secure elements needed to draw reasonable conclusions; to form conjectures and hypotheses; to consider relevant information and to deduce the consequences flowing from data, statements, principles, evidence, judgments, beliefs, opinions, concepts, descriptions, questions, or other forms of representation.' As sub-skills of inference the experts list querying evidence, conjecturing alternatives, and drawing conclusions.*

Beyond being able to interpret, analyze, evaluate and infer, strong critical thinkers can do two more things. They can explain what they think and how they arrived at that judgment. And, they can apply their powers of critical thinking to themselves and improve on their previous opinions. These wo skills are called 'explanation' and 'self-regulation.' The experts define ***explanation*** *as being able to present in a cogent and coherent way the results of one's reasoning. This means to be able to give someone a full look at the big picture: both 'to state and to justify that reasoning in terms of the evidential, conceptual, methodological, criteriological, and contextual considerations upon which one's results were based; and to present one's reasoning in the form of cogent arguments.' The sub-skills under explanation are describing methods and results, justifying procedures, proposing and defending with good reasons one's causal and conceptual explanations of events or points of view, and presenting full and well- reasoned, arguments in the context of seeking the best understandings possible. …*

Maybe the most remarkable cognitive skill of all, however, is this next one. This one is remarkable because it allows strong critical thinkers to improve their own thinking. In a sense this is critical thinking applied to itself. Because of that some people want to call this 'meta-cognition,' meaning it raises thinking to another level. But 'another level' really does not fully capture it, because at that next level up what self-regulation does is look back at all the dimensions of critical thinking and double check itself. … 'self-consciously to monitor one's cognitive activities, the elements used in those activities, and the results educed, particularly by applying skills in analysis, and evaluation to one's own inferential judgments with a view toward questioning, confirming, validating, or correcting either one's reasoning or one's results.' The two sub- skills here are self-examination and self-correction."

Facione 2011

[19] *"What kind of a person would be apt to use their critical thinking skills? The experts poetically describe such a person as having 'a critical spirit.' Having a critical spirit does not mean that the person is always negative and hypercritical of everyone and everything. The experts use the metaphorical phrase critical spirit in a positive sense. By it they mean 'a probing inquisitiveness, a keenness of mind, a zealous dedication to reason, and a hunger or eagerness for reliable information. The kind of person being described here is the kind that always wants to ask 'Why?' or 'How?' or 'What happens if?'. The one key difference, however, is that in fiction Sherlock always solves the mystery, while in the real world there is no guarantee. Critical thinking is about how*

you approach problems, questions, issues. It is the best way we know of to get to the truth. … The approaches to life and living which characterize critical thinking include:

** inquisitiveness with regard to a wide range of issues,*
** concern to become and remain well-informed,*
** alertness to opportunities to use critical thinking,*
** trust in the processes of reasoned inquiry,*
** self-confidence in one's own abilities to reason,*
** open-mindedness regarding divergent world-views,*
** flexibility in considering alternatives and opinions,*
** understanding of the opinions of other people,*
** fair-mindedness in appraising reasoning,*
** honesty in facing one's own biases, prejudices, stereotypes, or egocentric tendencies,*
** prudence in suspending, making or altering judgments,*
** willingness to reconsider and revise views where honest reflection suggests that change is warranted.*

What would someone be like who lacked those dispositions? It might be someone who does not care about much of anything, is not interested in the facts, prefers not to think, mistrusts reasoning as a way of finding things out or solving problems, holds his or her own reasoning abilities in low esteem, is close-minded, inflexible, insensitive, cannot understand what others think, is unfair when it comes to judging the quality of arguments, denies his or her own biases, jumps to conclusions or delays too long in making judgments, and never is willing to reconsider an opinion. … The experts went beyond approaches to life and living in general to emphasize that strong critical thinkers can also be described in terms of how they approach specific issues, questions, or problems. The experts said you would find these sorts of characteristics: approaches to life and living in general to emphasize that strong critical thinkers can also be described in terms of how they approach specific issues, questions, or problems. The experts said you would find these sorts of characteristics:

** clarity in stating the question or concern,*
** orderliness in working with complexity,*
** diligence in seeking relevant information,*
** reasonableness in selecting and applying criteria,*
** care in focusing attention on the concern at hand,*
** persistence though difficulties are encountered,*
** precision to the degree permitted by the subject and the circumstances."*

Facione 2011

[20] *"One system is more intuitive, reactive, quick and holistic. So as not to confuse things with the notions of thinking in popular culture, cognitive scientists often name this system, 'System 1.' The other (yes, you can guess its name) is more deliberative, reflective, computational and rule governed. You are right, it is called 'System 2'. In* ***System 1*** *thinking, one relies heavily on a number of heuristics (cognitive maneuvers), key situational characteristics, readily associated ideas, and vivid memories to arrive quickly and confidently at a judgment. System 1 thinking is particularly helpful in familiar situations when time is short and immediate action is required. While System 1 is functioning, another powerful system is also at work, that is, unless we shut it down by abusing alcohol or drugs, or with fear or indifference. Called '****System 2****,' this is our more reflective thinking system. It is useful for making judgments when you find yourself in unfamiliar situations and have more time to figure things out. It allows us to process abstract concepts, to deliberate, to plan ahead, to consider options carefully, to review and revise our work in the light of relevant guidelines or standards or rules of procedure. While System 2 decisions are also influenced by the correct or incorrect application of heuristic maneuvers, this is the system which relies on well-articulated reasons and more fully developed evidence. It is reasoning based on what we have learned through careful analysis, evaluation, explanation, and self-correction. This is the system which values intellectual honesty, analytically anticipating what happens next, maturity of judgment, fair-mindedness, elimination of biases, and truth-seeking. This is the system which we rely on to think carefully through complex, novel, high-stakes, and highly integrative problems … Educators urge us to improve our critical thinking skills and to reinforce our disposition to use those skills because that is perhaps the best way to develop and refine our System 2 reasoning. System 1 and System 2 are both believed to be vital decision-making tools when stakes are high and when uncertainty is an issue. Each of these two cognitive systems are believed to be capable of functioning to monitor and potentially override the other. This is one of the ways our species reduces the chance of making foolish, sub-optimal or even dangerous errors in judgment. Human thinking is far from perfect. Even a good thinker makes both System 1 and 2 errors. At times we misinterpret things, or we get*

our facts wrong, and we make mistakes as a result. But often our errors are directly related to the influences and misapplications of cognitive heuristics."

Facione 2011

[21] *"Any defensible conception must construe critical thinking in such a way as to capture most of what people have in mind when they claim that developing critical thinking is an important goal of education. That is to say, it must be true to the core meaning of the educator's basic concept of critical thinking. Should it fail in this regard, it is largely irrelevant to educators concerned with developing critical thinking. What, then, do critical thinking advocates generally have in mind when they talk about critical thinking? … thinking regarded as critical thinking must be directed toward some end or purpose, such as answering a question, making a decision, solving a problem, resolving an issue, devising a plan, or carrying out a project. Roughly speaking, thinking that serves such purposes can be characterized as thinking aimed at forming a judgement, i.e. making up one's mind about what to believe or do. … Not just any thinking aimed at deciding what to believe or do can count as critical thinking, however. If the thinking is sloppy, superficial, careless, rash or naive, most advocates of critical thinking would not agree that it is critical thinking. For example, someone who comes to believe on the basis of poor or irrelevant reasons, on the authority of someone whose credibility is questionable, or without attempting to assess the evidence relevant to the truth of the belief, would not usually be regarded as thinking critically. This suggests that thinking about what to believe or do must meet appropriate standards if it is to be regarded as critical thinking. Moreover, these standards cannot be met merely by accident or happenstance. If someone were inadvertently to fulfill relevant standards in their thinking, but had not intentionally attempted to fulfill them, they would not generally be regarded as having engaged in critical thinking. To be engaged in critical thinking one must be aware that there are such standards and must be striving to fulfill them. This is not to say of course, that a person engaged in thinking critically is necessarily able to state or verbally explicate the relevant standards. To summarize, critical thinking, as it is typically understood by educators, has at least these three features:*

* it is done for the purpose of making up one's mind about what to believe or do,
** the person engaging in the thinking is trying to fulfill standards of*
** the adequacy and accuracy appropriate to the thinking and the thinking fulfill the relevant standards to some threshold level."*

Bailin et al. 1999

[22] *"...this construal is too narrow, in that it fails to do justice to the fact that critical thinking very often takes place in the context of persons' thinking things through together by means of discussion and dialogue. Popper (1972: 148) has emphasized the importance of critical discussion in the advancement of science. We believe it plays an equally important role in most areas of inquiry and practice, including political and moral decision making. If we are correct in supposing that group deliberation is an important context for critical thinking, then the thinking appropriate to such contexts must be included in our conception of critical thinking. This means that, in addition to assessing intellectual products appropriately, critical thinking will include responding constructively to reasons and arguments given by others in the context of discussion. Responding constructively in such a context involves furthering the point or purpose of the critical discussion, while maintaining a social environment that enables all parties to the discussion to participate fully. Thus, good thinking in this context involves more than good reason-evaluation."*

Bailin et al. 1999

[23] *"Because critical thinking is, in our view, thinking in such a way as to fulfill relevant standards, it is the standards of good thinking that provide the criteria for determining what attributes are important for critical thinkers. If an attribute is required by persons in order to fulfill a standard of good thinking, or if it will significantly increase the chances that their thinking fulfill such standards, it can legitimately be regarded as an attribute that should be fostered in a critical thinker."*

Bailin et al. 1999

[24] *"It is fairly common to characterize the critical thinker by enumerating a list of skills or abilities and a list of attitudes or dispositions such a thinker must have. This kind of characterization is appealing, because obviously there are certain kinds of things critical thinkers must be able and inclined to do. However, adoption of such language, and in particular the use of the terms 'skills' and 'abilities', has the potential for causing confusion. The lists of critical thinking skills or abilities offered by critical thinking theorists are necessarily a list of things the critical thinker must be able to accomplish, for the only way we have of describing what one is able to do in*

thinking is in terms of the outcomes generated by the thinking. Thus, we say that the critical thinker must be able to do such things as judge the adequacy of reportive definitions, detect invalid arguments, etc. … Such lists imply nothing about the psychological states, capacities or processes that enable critical thinkers to have the requisite accomplishments, and nothing about the kinds of instructional procedures that are likely to be efficacious in bringing them about. Misunderstanding arises, however, when we begin to describe the items on such a list as 'critical thinking abilities or skills', that is, when we begin to talk about the ability to judge reportive definitions' or 'the skill of detecting invalid arguments.' Many educators interpret such ability and skill descriptions as descriptions of psychological processes, states or capacities, rather than as simply descriptions of what persons can accomplish."

Bailin et al. 1999

[25] *"To a considerable extent, the quality of thinking persons are able to do about a particular problem, issue or question is determined by what they know, or are able to find out, about it and about the context in which it must be resolved. Moreover, critical thinking always takes place in the context of (and against the backdrop of) already existing concepts, beliefs, values, and ways of acting. This context plays a very significant role in determining what will count as sensible or reasonable application of standards and principles of good thinking. Thus, the depth of knowledge, understanding and experience persons have in a particular area of study or practice is a significant determinant of the degree to which they are capable of thinking critically in that area. For example, standards for assessing the strength of inductive evidence for an empirical generalization cannot be sensibly or sensitively applied without knowing something about the nature of the phenomena covered by the generalization, including background theories concerning it and related phenomena. Similarly, thinking critically in deciding whether to accept or reject a moral judgement requires a clear understanding of the nature of the action or policy being judged, the context in which it is to be carried out, and the range of moral considerations relevant to the judgement. …*

Every area of intelligent human inquiry and practice, including science, art, law and morality, embodies within it practices of criticism by which proposed conclusions or ways of acting are tested, and previously accepted beliefs, practices and institutions are criticized and revised. Implicit in these practices are standards of critical assessment. It is these standards that critical thinkers must learn to use. They include not just rules of logic, but also standards of practical deliberation, standards of argumentation, standards used in developing plans of action, standards governing judgements made in the course of action (as in artistic and athletic performances), and standards governing inquiry and justification in specialized areas of study such as art, biology, history, literary criticism, mathematics, and technology. … Because verbal formulations of principles of critical thinking are abstracted from good critical practice, they typically do not tell a thinker all there is to know about the principles and how to apply them. For this reason they cannot be applied in a mechanical fashion. Rather, their abstractness gives them a vagueness that makes it necessary for the critical thinker to exercise judgement in interpreting them and determining what they require in any particular case. To acquire judgement of this sort, it is necessary to understand the practices of which the critical thinking principles are a part, and the point or purpose of these practices. It is also necessary to be acquainted with exemplars of the use of good judgement in applying the principles in a variety of contexts, because such examples provide the best indication of how the principle applies to particular cases. ...

Standards and principles of critical thinking are cultural artifacts that may be, and sometimes are, criticized and altered on the basis of our collective experience in using them. …

Although such standards are not readily revised, the appropriateness of any particular standard or the force it should have in a given context may always become a matter for critical reflection. …

One must also have certain commitments, attitudes or habits of mind that dispose him or her to use these resources to fulfill relevant standards and principles of good thinking. Passmore (1967: 197) aptly characterizes the possession of these character traits as having a 'critical spirit'. Moreover, as Siegel (1988: 9) points out, the critical thinker's tendency to fulfill the standards and principles of good thinking cannot be mindless or simply the result of habituation. Rather, it must be based on a recognition of the value of critical thinking, i. e. its importance in fostering true belief and responsible action."

Bailin et al. 1999

[26] *"Thus, teaching critical thinking is best conceptualized not as a matter of teaching isolated abilities and dispositions, but rather as furthering the initiation of students into complex critical practices that embody value-commitments and require the sensitive use of a variety of intellectual resources in the exercise of good*

judgement. Initiation of children into these practices begins long before they enter school. By the time they are in primary school they are already making and criticizing judgements and arguments of various sorts, though their arguments and criticisms may not be very good. The educator's task is to continue the student's initiation in a more discriminating and self-conscious way, such that good critical practice is encouraged and poor practice is abandoned. This involves not simply teaching students standards and concepts of which they were previously ignorant, but also getting them to appreciate the value of changing some of their previously established commitments and practices. Although the long-range educational project is to develop competence in thinking critically in a variety of areas, the attainment of this end is necessarily a gradual process that can begin in the earliest years in school. Teaching students how to appraise evidential arguments in history or chemistry may have to await secondary school or university, but primary students can begin to learn important commitments and habits of mind, such as thinking reasons and truth are important, respecting others in discussion, being open-minded, and being willing to look at issues from others' points of view. They can learn a variety of critical concepts, such as those necessary for distinguishing between definitions and empirical statements; they can learn a number of heuristics, such as asking for examples when the meaning of a term is unclear and reminding themselves to double-check claims before accepting them as fact; and they can learn principles, such as trying to think of alternatives when deciding what is the best thing to do. As they become more mature they can expand the range of intellectual resources they are able and willing to employ and improve the judgement with which they employ them."

Bailin et al. 1999

[27] *"Mystery meatloaf*
It's lunchtime. Five classmates have assembled in the cafeteria and are surveying the lunch choice. Present are Phil Gold, Nancy McGregor, Ravi Singh, Ahmed Ali, and Sophia Onassis.
Phil: Ah — mystery meatloaf — my favorite. I'll have a big piece with lots of gravy — and a double order of fries.
Nancy: Gross!
Phil: Yeah— I guess you're right—double fries is a bit much.
Nancy: No— I mean — MEAT? RED meat, yuck! How can you eat that stuff?
Phil: So what's the big deal?
Ravi: Haven't you heard ? Our Nancy's become a vegetarian.
Ahmed: No way! Not one of those granola-munching hippies?
Nancy: I just don't see how you can possibly bring yourself to eat another living creature. It's cruel… and inhumane!
Ravi: But animals eat other animals. It's just natural.
Ahmed: And besides, meat tastes so good. Just think of biting into a big, juicy, pink steak… mmmm.
Nancy: Ugh!
Phil: Everyone eats meat— at least all normal people. It's just some dumb cow.
Sophia: I've heard that animals used for meat are kept in horrible conditions.
Ravi: So now you believe everything you hear? That's not like you, Sophia. It's all a load of propaganda from those animal rights loonies.
Nancy: You guys are just a bunch of… cannibals!
Ahmed: Well, isn't she on her high horse, dictating to us what we should and shouldn't eat!
Nancy: Come on, Sophia — let's move to another table. I can't sit with the… insensitive boors!
Phil: Fine! Now at least we can eat in peace. Bring on the meatloaf…
You have likely come across scenes like this, possibly quite often. They are, unfortunately, fairly typical. There is a disagreement over an issue about which people feel strongly. The disagreement escalates into name-calling and high emotions and ends in misunderstanding.
But this conversation could go in a very different direction.

Mystery Meatloaf – Take II
Phil: Ah — mystery meatloaf — my favorite. I'll have a big piece with lots of gravy – and a double order of fries.
Nancy: I'll have the vegetarian lasagna please, with a side of yam fries.
Ahmed: What, no meatloaf today?
Ravi: Haven't you heard? Our Nancy's become a vegetarian.
Ahmed: No way! Why did you do something like that?
Nancy: It just finally got to me that I was eating an animal, another living creature, and that didn't seem right.

Ravi: But animals eat other animals. It's just natural.
Ahmed: And besides, it tastes so good.
Phil: Anyway, it's just a dumb cow, isn't it? It doesn't have thoughts or feelings like a person, does it?
Ravi: I'm pretty sure that animals feel pain. My dog sure howls when he gets his tail caught in the door.
Phil: Well, what about fish? They're not too with it.
Nancy: Some of my vegetarian friends do eat fish. I've been struggling with that one.
Sophia: I've heard that animals used for meat are kept in horrible conditions.
Ravi: I wonder if that's true or whether it's mostly propaganda from the animal rights folks?
Sophia: I haven't really checked it out …
Ahmed: And there are some animals that live quite well. There are those free-range chickens who get to roam around and have lots of grain to eat and lead a normal chicken life (in fact, probably better than most). Until it's time to hop into the pot, that is. So is it OK to eat those free-range chicks?
Nancy: But it's still killing other living creatures for our selfish purposes. Why should we think that human beings have a right to do that?
Phil: It does bother me, though, when folks get so worked up about how we treat animals, especially cute ones with big eyes, and ignore all the people getting mistreated and even killed all over the world. Isn't that more important?
Sophia: Like the way all those movie stars and famous people protest about the seal hunt in Newfoundland but don't take any action about all the genocides happening around the world.
Ahmed: Wow – we've sure come up with a lot of questions. Though not many answers.
Sophia: I wonder ... maybe there's some way to go about trying to answer some of the questions. We couldn't be the first people to think about these issues. So we could have a look to see what ideas and information are out there.
Nancy: I'm sure there's information about the conditions in which animals are kept.
Ravi: And there must be research about whether different animals can feel pain, or even have other feelings.
Phil: And I'll bet other folks have thought about the moral issues about the treatment and rights of animals. I'd be interested in seeing what's been written about that.
Ravi: Though I don't expect that they'll all agree.
Phil: No, but that would at least give us some ideas to consider.
Ravi: And evaluate.
Sophia: I think it's worth a try. I don't know if we'll end up agreeing. Maybe. But even if we don't, at least we'll be able to think about the issues in a more informed way. And we'll be able to understand where the others are coming from.
Nancy: Now that would be progress!"

Bailin & Battersby 2016.

[28] "*Key aspects of critical thinking, as currently advocated by contemporary theorists, includes: 1. the claim that the notion is essentially normative in character; and 2. the claim that critical thinking involves two distinct components: a. skills or abilities of reason assessment and b. the dispositions to engage in and be guided by such assessments. …*

So understood, critical thinking is a sort of good thinking. Therefore the notion of critical thinking is fundamentally a normative one, thus distinguishing this understanding of critical thinking from those, common in psychology, which treat the notion as descriptive, identifying particular psychological processes (Bailin et al 1999). To characterize a given episode of thinking as critical is to judge that it meets relevant standards or criteria of acceptability, and is thus appropriately thought of as good. …

Robert H. Ennis, for example, defines critical thinking as 'reasonable reflective thinking that is focused on deciding what to believe and do.' (Ennis, 1987), and offers a detailed list of abilities, skills and dispositions which thinking (and thinkers) must manifest if it is (they are) to qualify as critical. Matthew Lipman defines critical thinking as thinking that facilitates judgment because it relies on criteria, is self-correcting and is sensitive to context (Lipman 1991). Richard Paul analyses critical thinking in terms of the ability and the disposition to critically evaluate beliefs, their underlying assumptions, and the worldviews in which they are embedded (Paul 1990). Harvey Siegel characterizes the critical thinker as one who is 'appropriately moved by reasons' (Siegel 1988), and emphasizes the critical thinker's mastery of epistemic criteria which reasons must meet in order to be rightly judged to be good reasons, that is, reasons that justify beliefs, claims, judgments, and actions. Other authors, including John McPeck (1981, 1990) similarly emphasize the normative dimension of the

concept. While these authors' accounts of critical thinking differ on many respects, and have their own emphases, they are nevertheless agreed on its essentially normative character (Bailin & Siegel 2003). … While some early treatments of critical thinking defined it only in terms of particular skills, … almost all recent philosophical discussion of it… regards critical thinking as involving both 1. skills or abilities of reason assessment and 2. a cluster of dispositions, habits of mind, and character traits, sometimes referred to collectively as the critical spirit (Siegel 1988)."

Siegel 2010

[29] "*The ideal calls for the fostering of certain skills and abilities, and for the fostering of a certain form of character. it is thus a general ideal of a certain sort of person, which sort of person it is the task of education to help to create. This aspects of the educational ideal of rationality aligns it with the complementary ideal of autonomy, since a rational person will – at least ideally – also be an autonomous one, capable of judging for himself/herself the justifiedness of candidate beliefs and the legitimacy of candidate values."*

Siegel 2010

[30] "*Ennis (1989) identifies a range of assumptions regarding domain specificity held by various theorists. For example, most researchers view background knowledge as a necessary but not sufficient condition for critical thinking. In addition, some researchers see the transfer of critical thinking skills across domains as unlikely unless students are provided with sufficient opportunities to practice these skills in a variety of domains and the students are explicitly taught to transfer. Finally, an even smaller number of researchers hold the view that general instruction in critical thinking skills is unlikely to be successful because critical thinking skills are inherently domain-specific (Ennis, 1989). Proponents of domain specificity include Willingham (2007), who argues that it is easier to learn to think critically within a given domain than it is to learn to think critically in a generic sense. Similarly, Bailin (2002) argues that domain-specific knowledge is necessary for critical thinking because what constitutes valid evidence, arguments, and standards tends to vary across domains. …*

Although McPeck (1990) concedes that there are a limited number of general thinking skills, he argues that the most useful thinking skills are those that are domain-specific. According to McPeck, the more general the thinking skill, the less helpful it is. …

Those who maintain that critical thinking skills and abilities are not domain-specific include Halpern (2001), who reviews evidence on the success of general instruction in critical thinking skills and concludes that such instruction has great potential. Lipman (1988) notes that critical thinking facilitates good judgment because it relies on criteria. These criteria may differ across domains, but the fundamental meaning of critical thinking remains the same. Van Gelder (2005, p. 43) argues that critical thinking is 'intrinsically general in nature,' which, paradoxically, is why critical thinking skills and abilities are notoriously difficult to transfer to new contexts."

Lai 2011

[31] "*Although the ability to think critically has always been important, it is a vital necessity for the citizens of the 21st century. …*

The workforce is one critical place where we can witness the dizzying pace of change. There is an increased demand for a new type ion worker – the knowledge worker or the symbol analyst – …. The information explosion is yet another reason why we need to provide specific instruction in thinking. … Relevant, credible information has to be selected, interpreted, digested, evaluated, learned, and applied … The twin abilities of knowing how to learn and knowing how to think clearly about the rapidly proliferating information that we must select from are the most important intellectual skills of the 21st century. …

For the first time in history, we have the ability to destroy all life on earth. The decisions that we make as individuals and as a society regarding the economy, conservation of the natural resources, and the development of nuclear weapons will affect future generations of all people around the world."

Halpern 2013

[32] "*Critical thinking refers to the use of cognitive skills or strategies that increase the probability of a desirable outcome. Critical thinking is purposeful, reasoned, and goal-directed. It is the kind of thinking involved in solving problems, formulating inferences, calculating likelihoods, and making decisions. Critical thinkers use these skills appropriately, without prompting, and usually with conscious intent, in a variety of settings. That is, they are predisposed to think critically. When we think critically, we are evaluating the outcomes of our thought*

processes—how good a decision is or how well a problem is solved (Halpern, 1996, 1998). This definition is broad enough to encompass a variety of viewpoints, so critical thinking can be taught as argument analysis (see, for example, Kahane, 1997), problem solving (Mayer, 1992), decision making (Dawes, 1988), or cognitive process (Rabinowitz, 1993). Regardless of the academic background of the instructor or the language used to describe critical thinking, all of these approaches share a set of common assumptions: there are identifiable critical thinking skills that can be taught and learned, and when students learn these skills and apply them appropriately, they become better thinkers."

Halpern 1999

[33] *"What's 'critical' about critical thinking? The 'critical' part of critical thinking denotes an evaluation component. Sometimes the word 'critical' is used to convey something negative, as when we say 'she is a critical person'. However, evaluation can and should be a constructive reflection of positive and negative attributes. When we think critically, we are evaluating the outcomes of our thought processes – how good a decision is and how well a problem has been solved. Critical thinking also involves evaluating the thinking process – the reasoning that went into the conclusion we've arrived at or the kinds of factors considered in making g a decision."*

Halpern 2013

[34] "*Often there is noncritical, or more appropriately labelled, rote memorization or lower level thinking that is taught and tested in many classrooms, at all levels of education at the expense of higher ordered critical thinking. … Knowledge about a content area is critical to critical thinking ; no one can think critically about any topic without the necessary background information, but the facts alone are not enough. …*

Critical thinking skills are often referred to as 'higher order thinking skills' to differentiate them from simpler (i.e., lower order) thinking skills. Higher order skills are relatively complex; require judgment, analysis, and synthesis; and are to applied in rote and mechanical manner. Higher order thinking is thinking that is reflective, sensitive to the context, and self-monitored. Computational arithmetic, for example, is not a higher order skill, even though it is an important skill, because it involves the rote application of well-learned rules with little concern for context or other variables that would affect the outcome. By contrast, deciding which of two information sources is more credible is a higher order cognitive skill because it is a judgment task in which the variables that affect credibility are multidimensional and change with the context. In real life critical thinking skills are needed whenever we grapple with complex issues and messy, ill-defined problems."

Halpern 2007

[35] "*A general list of skills that would be applicable in almost any class would include understanding how cause is determined, recognizing and criticizing assumptions, analyzing means-goals relationships, giving reasons to support a conclusion, assessing degrees of likelihood and uncertainty, incorporating isolated data into a wider framework, and using an alogies to solve problems."*

Halpern 1998

[36] "*A short taxonomy of critical-thinking skills is proposed as a guide for instruction: (a) verbal reasoning skills--This category includes those skills needed to comprehend and defend against the persuasive techniques that are embedded in everyday language; (b) argument analysis skills--An argument is a set of statements with at least one conclusion and one reason that supports the conclusion. In real-life settings, arguments are complex, with reasons that run counter to the conclusion, stated and unstated assumptions, irrelevant information, and intermediate steps; (c) skills in thinking as hypothesis testing--The rationale for this category is that people function like intuitive scientists to explain, predict, and control events. These skills include generalizability, recognition of the need for an adequately large sample size, accurate assessment, and validity, among others; (d) likelihood and uncertainty--Because very few events in life can be known with certainty, the correct use of cumulative, exclusive, and contingent probabilities should play a critical role in almost every decision; (e) decision-making and problem-solving skills--In some sense, all of the critical-thinking skills are used to make decisions and solve problems, but the ones that are included here involve generating and selecting alternatives and judging among them. Creative thinking is subsumed under this category because of its importance in generating alternatives and restating problems and goals."*

Halpern 1998

[37] "*I am sometimes told that there is no such thing as critical thinking because different viewpoints are all a matter of opinion and that everyone has a right to his or her own opinion. They argue that a better way to think*

does not exist. I certainly agree that we all have the right to our own opinion; however some opinions are better than others. If, for example you believe that heavy alcohol consumption is good for pregnant women, you had better to back up this belief with a sound reasoning. …

Similarly, everyone has the right to believe in phenomena such as astrology and extra-sensory perception, but there is no sound evidence to support the existence of these phenomena."

Halpern 2013

[38] "*Many authorities in higher education did not enthusiastically embrace the idea that college students should receive explicit instruction in how to think. Not that the academic community was opposed to good thinking, but many educators believed that it was a misguided effort. For example, Glaser (1984) cited abundant evidence of critical thinking failures in support of his argument that thinking skills are context-bound and do not transfer across academic domains. Glaser and other skeptics were partly correct. Better thinking is not a necessary outcome of traditional, discipline-based instruction. However, when thinking skills are explicitly taught for transfer, using multiple examples from several disciplines, students can learn to improve how they think in ways that transfer across academic domains."*

Halpern 1999

"We now have a considerable body of evidence that thinking skills courses and thinking skills instruction that is embedded in other courses can have positive effects that are transferable to many situations. …

The whole enterprise of learning how to improve thinking would be of little value if these skills were used only in the classroom or only on problems that are very similar to those presented in class. Ideally, critical thinking skills should be used to recognise and resist unrealistic campaign promises, circular reasoning, faulty probability estimates, weak arguments by analogy, or language designed to mislead whenever and wherever it is encountered. Critical thinkers should be able to solve or offer reasonable solutions to real-world problems, whether to solve the problem of nuclear war or how to set up a new video recorder."

Halpern 2013

[39] "*For example, students may be able to explain why correlation is not causation when presented with this question on an exam but still not recognize that this same principle is operating when they read that children who attend religious schools score higher on standardized tests than those who attend public schools. Specific instruction in recognizing the structure of correlational problems can improve the probability that students will recognize these problems, even when the topic is different."*

Halpern 1999

[40] "*It is important to separate the disposition or willingness to think critically from the ability to think critically. Some people may have excellent critical-thinking skills and may recognize when the skills are needed, but they also may choose not to engage in the effortful process of using them. This is the distinction between what people can do and what they actually do in real-world contexts. It is of no value to teach students the skills of critical thinking if they do not use them. Good instructional programs help learners decide when to make the necessary mental investment in critical thinking and when a problem or argument is not worth the effort. An extended session of generating alternatives and calculating probabilities is a reasonable response to a diagnosis of cancer; it is not worth the effort when the decision involves the selection of an ice-cream flavor."*

Halpern 2013

[41] "*Critical thinking is effortful: it requires a concern for accuracy and the willingness to persist at difficult tasks Adeline suppress immediate Adeline easy responses. It requires an openness to new ideas, which some people find to be the most difficult component."*

Halpern 1999

[42] "*For example, making judgments about the likelihood of remembering a fact or event at some time in the future, or deciding how well a problem has been solved, or estimating one's own performance on a test of comprehension of complex prose. The underlying idea is that everyone needs to be able to assess how well they are thinking or how much they know ... about a topic to make reasoned decisions. Research has shown that when people have little knowledge of a content area (e.g. logical reasoning) they will misperceive their ability and take themselves to be much more higher in ability than they actually are …*

Students can become better thinkers and learners by developing the habit of monitoring their understanding and by judging the quality of their learning. …

Instruction to enhance the development of critical cal thinking skills should include a metacognitive component."

Halpern 1999

[43] *"Learning tasks, like real-world thinking tasks, should be rich in information. Some of the information available may not be relevant, and part of the learning exercise involves deciding which information is important to the problem. What is important in the teaching and learning of critical-thinking skills is what the learners are required to do with the information. Learning exercises should focus on the critical aspects of the problems and arguments that utilize the skills. The tasks should require thoughtful analysis and synthesis. For example, the repeated use of 'authentic' materials, or materials that are similar to real-world situations, is one teaching strategy to enhance transfer (Derry, Levin, & Schauble, 1995). Thinking skills need to be explicitly and consciously taught and then used with many types of examples so that the skill aspect and its appropriate use are clarified and emphasized."*

Halpern 1999

[44] *"1. Explicitly learn the skills of critical thinking*
2. Develop the dispositions for effortful thinking and learning
3. Direct learning activities in ways that increase the probability of trans contextual transfer (structure training)
4. Make metacognitive monitoring explicit and overt."

Halpern 2013

[45] *"Most instructional programs designed to teach critical thinking do not draw on contemporary empirical research in cognitive development as a potential resource. The developmental model of critical thinking outlined here derives from contemporary empirical research on directions and processes of intellectual development in children and adolescents. It identifies three forms of second-order cognition (meta-knowing) - metacognitive, metastrategic, and epistemological - that constitute an essential part of what develops cognitively to make critical thinking possible."*

Kuhn 1999

"... the developing cognitive competencies I describe as most relevant to critical thinking are metacognitive - rather than cognitive-competencies. In contrast to first-order cognitive skills that enable one to know about the world, metacognitive skills are second-order meta-knowing skills that entail knowing about one's own (and others') knowing. …

It should be noted, however, that a concept like metacognition, even if it does not go by this specific name, is by no means new to the philosophical literature on critical thinking. Indeed, something like metacognition figures in the definitions of critical thinking proposed by most educational philosophers who have addressed the topic. …

In contrast, an aspect of critical thinking that has received relatively little attention from educational philosophers is its developmental dimension. One important exception to this generalization, however, is the work of the educational philosopher who perhaps had the most to teach us about critical thinking - Dewey. Dewey, with his deep respect for and involvement in the empirical science of psychology, did not share the conception that prevails today of a dichotomy between the scientific study of human development and the practical world of children in schools. Instead, he saw schools as laboratories of human development, as experiments in the possibilities of human development in arranged environments. Repeatedly in his writings, Dewey made clear that the goal of education could not only be development (or what he called 'growth'). Education 'means supplying the conditions which foster growth' (Dewey, 1916, p. 56), not toward a predetermined end but rather in the direction of 'an increase in the range and complexity of situations to which the child is capable of applying reasoned inquiry' (Cahan, 1994, p. 158). Dewey also made it clear that he saw the educator's task as a process of connecting with the young child's interests and purposes, but that one could not stop there. 'The real problem of intellectual education,' he said, 'is the transformation of more or less casual curiosity and sporadic suggestion into attitudes of alert, cautious, and thorough inquiry' (Dewey, 1933, p. 181)."

Kuhn 1999

[46] *"A second distinctive characteristics of the present effort is that the developing cognitive competencies I describe as most relevant to critical thinking are metacognitive - rather than cognitive - competencies. In contrast to first-order cognitive skills that enable one to know about the world, metacognitive skills are second order meta-knowing skills that entail knowing about one's own (and others') knowing."*

Kuhn 1999

"The development of metacognitive understanding is essential to critical thinking because critical thinking by definition involves reflecting on what is known and how that knowledge is justified. Individuals with well-developed metacognitive skills are in control of their own beliefs in the sense of exercising conscious control over their evolution the face of external influences. They know what they think and ca justify why. Their skills in the conscious coordination of theory and evidence also put them in a position to evaluate the assertions of others."

Kuhn 1999

[47] *"Like pre-schoolers many older individuals blur the distinction between theory-based and evidence-based sources of their beliefs. Rather than seeing their theories as belief states subject to disconfirmation and representing theory and evidence as distinct entities to be reconciled with one another, they merge the two into a single representation of 'the way things are' with little apparent awareness of the source of their belief. Evidence serves merely to illustrate what one knows to be true…*

Theories may eventually change in response to discrepant evidence, but often with the individual manifesting little awareness or control of the process. Like young children… older participants in our studies are likely to deny that they ever held a belief different from the one they are now professing."

Kuhn 1999

[48] *"Each of the three kinds of meta-knowing that have been examined here-the metacognitive, metastrategic, and epistemological - is central to critical thinking. The development of metacognitive understanding is essential to critical thinking because critical thinking by definition involves reflecting on what is known and how that knowledge is justified. Individuals with well-developed metacognitive skills are in control of their own beliefs in the sense of exercising conscious control over their evolution in the face of external influences. They know what they think and can justify why. Their skills in the conscious coordination of theory and evidence also put them in a position to evaluate the assertions of others. Metastrategic skill is also essential to critical thinking. Those who have developed strong metastrategic skills apply consistent standards of evaluation across time and situations. They do not succumb to a view of a favored assertion as more probable than its alternatives because of its favored status and, therefore, subject to different standards of evolution. They also resist the temptation of local interpretation (Klahr, Fay, & Dunbar, 1993) of an isolated piece of evidence as supportive because it is considered out of the context of a broader pattern of which it is a part. The development of epistemological understanding may be the most fundamental underpinning of critical thinking. If knowledge is entirely objective, certain, and simply accumulates, unconnected to the human minds that do this knowing-as the absolutist conceives-or if knowledge is entirely subjective, subject only to the tastes and wishes of the knower-as the multiplist conceives-critical thinking and judgment are superfluous. People must see the point of thinking if they are to engage in it. Put simply by one of the multiplists in our studies, 'I feel it's not worth it to argue because everyone has their opinion.' In such cases, educators can undertake to teach intellectual skills, but the reasons to apply them will be missing."*

Kuhn 1999

[49] *"In recent decades, philosophers, researchers and educators have focused their attention on a form of reasoning that combines reflection, justification, and the application and adaptation of individual reasoning; this has come to be known as critical thinking (Kuhn & Dean, 2004)."*

Kuhn 2018

[50] *"Each thirty-minute lesson was conducted in a classroom once every three days. The seminar included explicit instruction in regards to the development and analysis of arguments, counter-arguments and the value of trying to understand various views. It was designed to offer a nurturing environment and encouraged discussions regarding metacognitive practice. Students were asked to think about why they believe what they believe, and how they came to hold their beliefs. Discussions focused on understanding and analyzing the strength of reasons and arguments. The scrutiny of various claims and beliefs, along with an emphasis on open mindedness in the pursuit of good reasoning, were central themes of the seminar. Following the approach described by Kuhn and*

Udell (2003), the sessions included a great deal of dialogue and debate between groups of students who, through explicit instruction, had developed arguments for group discussion and analysis; together, students with opposing views discussed and analyzed their co-created arguments."

Kuhn 2018

"The interview protocol contained six open-ended questions that focused on the following characteristics: judging the credibility of the source; identifying conclusions, reasons and assumptions (evidence); judging the quality of an argument (cogency); developing a position; open-mindedness (considering counter-arguments), and metacognitive awareness."

Kuhn 1992

[51] *"Twentieth-century psychologists have been pessimistic about teaching reasoning, prevailing opinion suggesting that people may possess only domain-specific rules, rather than abstract rules; this would mean that training a rule in one domain would not produce generalization to other domains. Alternatively, it was thought that people might possess abstract rules (such as logical ones) but that these are induced developmentally through self-discovery methods and cannot be trained. Research suggests a much more optimistic view: even brief formal training in inferential rules may enhance their use for reasoning about everyday life events. Previous theorists may have been mistaken about trainability, in part because they misidentified the kind of rules that people use naturally."*

Nisbett et al. 1987

[52] "*Our initial work on the use and trainability of inferential rules focused on a set of statistical rules that are derivable from the law of large numbers. We and our colleagues have found that people reason in accordance with the law of large numbers in a wide range of tasks and domains. For example, generalization often proceeds in accordance with the principle that larger samples are required when generalizing about populations that are more variable with respect to a given attribute than when generalizing about populations that are less variable. Three results support the view that people possess an abstract version of the law of large numbers and that improvements to it can transfer to a wide range of problem content. First, purely abstract rule training produced improvement in both the frequency and the quality of statistical answers. Second, the abstract rule training effect was substantial across all three problem domains: training improved statistical reasoning for problems that people rarely think of in terms of probability just as much as it did for problems that people almost always think of in probabilistic terms. Since highly abstract statistical rules can be taught in such a way that they can be applied to a great range of everyday life events, is the same true of the even more abstract rules of deductive logic? We can report no evidence indicating that this is true, and we can provide some evidence indicating that it is not."*

Nisbett et al.1987

[53] "*Inferences frequently violate principles of statistics, economics, logic, and basic scientific methodology."*

Nisbett 2015

[54] "*Initially I was quite dubious that a course or two dealing with one or another approach to reasoning could have the kind of impact on people that long exposure to the concepts had on me. The XXth century skepticism about the possibility of teaching reasoning continued to influence my thinking. I could not have been more mistaken. It turns out that the courses people take in college do affect inferences about the world - often very markedly. Rules of logic, statistical principles such as how large numbers and regression to the mean, principles of scientific methodology such as how to establish control groups when making assertions about cause and effect, classical economic principles, and decision theory concepts all influence the way people think about problems that crop up in everyday life. The affect how people reason about athletic events, what procedures they think are best for going about the process of hiring someone, and event their approach to such minor questions as whether they should finish a meal that isn't very tasty. ...*

Most impressively, we sometimes questioned subjects weeks later, in contexts where they didn't know they were being studied, such as telephone polls allegedly being conducted by survey researchers. We were delighted to find that people often retained substantial ability to apply the concepts to ordinary problems outside the laboratory contexts in which the concepts had been taught."

Nisbett 2015

[55] *"Remember that all perceptions, judgments, and beliefs are inferences and are not direct readouts of reality. This recognition should prompt an appropriate humility about just how certain we should be about our judgments, as well as recognition that the views of other people that differ from our own may have more validity than our intuitions tell us they do.*
Remember that incidental, irrelevant perceptions and cognitions can affect our judgment and behavior.
Be alert of the possible role of heuristics in producing judgments.
Pay more attention to the context.
These injunctions can become part of the mental equipment you use to understand the world."

Nisbett 2015

[56] *"We don't normally think of forming impressions of individual's personality as a statistical process consisting of sampling a population of events, but they are exactly that. and framing them in that way makes us both more cautious about some kinds of personality ascriptions and better able to predict individual's behavior in the future. ...*

(1) frame everyday life events in such a way that the relevance of statistical principles is obvious and you can make contact with them, and (2) code the events in such a way that approximate versions of statistical rules can be applied to them… Once you have the knack of framing real-world problems as statistical ones and coding their elements in such a way that statistical heuristics can be applied, those principles seem to pop up magically to help you solve a given problem - often without conscious awareness that you're applying a rough-and-ready version of a statistical principle."

Nisbett 2015

"What counts as an explanation (for everything from why your friend acts in such an annoying way to why a product launch failed)? How can we tell the difference between events that are causally related and events that are merely associated to one another in time or place? What kinds of knowledge can be considered certain and what kinds only conjectural? What are the characteristics of a good theory - in science and in everyday life? How can we tell the difference between theories that can be falsified and theories that cannot?"

Nisbett 2015

[57] *"The distinction between two kinds of thinking, one fast and intuitive, the other slow and deliberative, is both ancient in origin and widespread in philosophical and psychological writing. Such a distinction has been made by many authors in many fields, often in ignorance of the related writing of others (Frankish & Evans, 2009). Our particular interest is in dual-process accounts of human reasoning and related higher cognitive processes, such as judgment and decision making. Such theories have their origins in the 1970s and 1980s (Evans, 1989; Wason & Evans, 1975) and have become the focus of much interest in contemporary research on these topics (Barbey & Sloman, 2007; Evans, 2007a, 2008; Evans & Over, 1996; Kahneman, 2011; Kahneman & Frederick, 2002; S. A. Sloman, 1996; Stanovich, 1999, 2011; Stanovich & West, 2000). Over a similar period, dual-process theories have proved popular in the psychology of learning (e.g., Dienes & Perner, 1999; Reber, 1993; Sun, Slusarz, & Terry, 2005) and especially in social cognition, which has the greatest proliferation of dual-processing labels and theories (see Chaiken & Trope, 1999; Epstein, 1994; Kruglanski & Orehek, 2007; Smith & DeCoster, 2000). Originally, dual-process theories in these different fields developed independently, although there have more recently been attempts to connect them. One consequence has been the development of broad dual-system theories that attempt to link a wide range of attributes to two systems of thought that are believed to underlie intuitive and reflective processing, respectively (Epstein, 1994; Evans & Over, 1996; Reber, 1993; Stanovich, 1999). Following Stanovich (1999), these are often referred to as Systems 1 and 2. As the popularity of dual-process and dual-system theories has increased, so too have the voices of criticism, as illustrated in the opening quotations. Critics have pointed to the multitude of dual-processing accounts, the vagueness of their definition, and the lack of coherence and consistency in the proposed cluster of attributes for two-system accounts. They have questioned the evidence on which such claims are made and have argued that single-process accounts can explain the data (Gigerenzer & Regier, 1996; Keren & Schul, 2009; Kruglanski et al., 2003; Kruglanski & Gigerenzer, 2011; Osman, 2004). Here we collaborate for the first time to respond to these various critiques. It is important that we do so, as although a number of these criticisms have some force to them (and have been acknowledged in our own recent writing), we believe that the dual-processing distinction is nonetheless strongly supported by a wide range of converging experimental, psychometric, and neuroscientific methods. In general, these critiques are problematic because they attack not any particular theory but rather a class of theories, effectively treating all dual-process and dual- system theories alike. However, all dual-process theories are not, by any means, the same. Our own work has developed dual-process theories of reasoning and*

decision making, but even in this domain, there is much in the writings of other authors with which we have disagreements."

Evans & Stanovich 2013

[58] "*Modes of processing—more commonly termed thinking dispositions—are well represented in Stanovich's (2009a, 2009b, 2011) tripartite model of mind displayed in its simplest form in Figure 1. In the spirit of Dennett's (1996) book Kinds of Minds, the set of autonomous systems (the source of Type 1 processing) is labeled as the autonomous mind, the algorithmic level of Type 2 processing the algorithmic mind, and the reflective level of Type 2 processing the reflective mind. Dennett's 'kinds of minds' terminology refers to hierarchies of control rather than separate systems. Two levels of control are associated with Type 2 processing and one with Type 1 processing. The autonomous set of systems (TASS) will implement their short-leashed goals unless overridden by an inhibitory mechanism of the algorithmic mind. But override itself is initiated by higher level control. That is, the algorithmic level is subordinate to higher level goal states and epistemic thinking dispositions. These goal states and epistemic dispositions exist at what might be termed the reflective level of processing—a level containing control states that regulate behavior at a high level of generality. Such high-level goal states are common in the intelligent agents built by artificial intelligence researchers (A. Sloman & Chrisley, 2003). The difference between the algorithmic mind and the reflective mind is captured in the well-established distinction in the measurement of individual differences between cognitive ability and thinking dispositions (and represented in Fig. 1). The former are measures of the ability of the algorithmic mind to sustain decoupled representations (for purposes of inhibition or simulation, see Stanovich, 2011). In contrast, thinking dispositions are measures of the higher level regulatory states of the reflective mind: the tendency to collect information before making up one's mind, the tendency to seek various points of view before coming to a conclusion, the disposition to think extensively about a problem before responding, the tendency to calibrate the degree of strength of one's opinion to the degree of evidence available, the tendency to think about future consequences before taking action, and the tendency to explicitly weigh pluses and minuses of situations before making a decision."*

Evans & Stanovich 2013

[59] "*Critical thinking is highly valued in educational writings if not in practice. Despite a substantial literature on the subject, for many years the area of critical thinking was notorious for its conceptual difficulties. For example, years ago Cuban (1984) lamented that 'defining thinking skills, reasoning, critical thought, and problem solving is troublesome to both social scientists and practitioners. Troublesome is a polite word; the area is a conceptual swamp' (p. 676). There has been some progress in elucidating the concept of critical thinking since the time of Cuban's statement, but we shall argue here that educational theory is on the verge of an even more stunning conceptual advance in the area of critical thinking. Education is beginning to understand the critical thinking concept by relating it to the constructs of intelligence and rational thought. In fact, modern cognitive science provides a coherent framework for understanding the relation between critical thinking, intelligence, and rational thought."*

Evans & Stanovich 2010

[60] "*It can be argued that the superordinate goal we are actually trying to foster is that of rationality (Stanovich, 2004, 2009). That is, much of what educators are ultimately concerned about is rational thought in both the epistemic sense and the practical sense. We value certain thinking dispositions because we think that they will at least aid in bringing belief in line with the world and in achieving our goals. …*

We can see that it is rationality, and not critical thinking per se, that is the ultimate goal of education by conducting some simple thought experiments or imaginative hypotheticals. For example, we could imagine a person with excellent epistemic rationality (his or her degree of confidence in propositions being well calibrated to the available evidence relevant to the proposition) and optimal practical rationality (the person optimally satisfies desires) who was not actively open-minded— that is, who was not a good critical thinker under standard assumptions. Of course we would still want to mold such an individual's dispositions in the direction of open-mindedness for the sake of society as a whole. But the essential point for the present discussion is that, from a purely individual perspective, we would now be hard pressed to find reasons why we would want to change such a person's thinking dispositions —whatever they were—if they had led to rational thought and action in the past. …

In short, a large part of the rationale for educational interventions to change thinking dispositions derives from a tacit assumption that actively open-minded critical-thinking dispositions make the individual a more rational

person—or as Sternberg (2001, 2005) argues, a wiser, less foolish person. Thus, the normative justification for fostering critical thought is that it is the foundation of rational thought. The thinking dispositions associated with critical thinking must be fostered because they make students more rational. Our view is consistent with that of many other theorists who have moved toward conceptualizing critical thinking as a subspecies of rational thinking or at least as closely related to rational thinking (Kuhn, 2005; Moshman, 2004, 2005, 2010; Reyna, 2004; Siegel, 1988, 1997). …

The grounding of critical thinking within the concept of rationality in this manner has many conceptual advantages. First, the concept of rationality is deeply intertwined with the data and theory of modern cognitive science (see LeBoeuf & Shafir, 2005; Over, 2004; Samuels & Stich, 2004; Stanovich, 2004, 2009) in a way that the concept of critical thinking is not."

Evans & Stanovich 2010

[61] "*Cognitive scientists recognize two types of rationality: instrumental and epistemic. The simplest definition of instrumental rationality is behaving in the world so that you get exactly what you most want, given the resources (physical and mental) available to you. Somewhat more technically, we could characterize instrumental rationality as the optimization of the individual's goal fulfillment. Economists and cognitive scientists have refined the notion of optimization of goal fulfillment into the technical notion of expected utility. The model of rational judgment used by decision scientists is one in which a person chooses options based on which option has the largest expected utility (see Baron, 2008; Dawes, 1998; Hastie & Dawes, 2001; Wu, Zhang, & Gonzalez, 2004). The other aspect of rationality studied by cognitive scientists is termed epistemic rationality. This aspect of rationality concerns how well beliefs map onto the actual structure of the world. Epistemic rationality is sometimes called theoretical rationality or evidential rationality (see Audi, 1993, 2001; Foley, 1987; Harman, 1995; Manktelow, 2004; Over, 2004). Instrumental and epistemic rationality are related. The aspect of beliefs that enter into instrumental calculations (that is, tacit calculations) are the probabilities of states of affairs in the world."*

Evans & Stanovich 2010

[62] "*Researchers in the tradition of heuristics and biases have demonstrated in a host of empirical studies that people violate many of the strictures of rationality and that the magnitude of these violations can be measured experimentally. For example, people display confirmation bias, test hypotheses inefficiently, display preference inconsistencies, do not properly calibrate degrees of belief, overproject their own opinions onto others, combine probabilities incoherently, and allow prior knowledge to become implicated in deductive reasoning (for summaries of the large literature, see Baron, 2008; Evans, 1989, 2007; Gilovich, Griffin, & Kahneman, 2002; Kahneman & Tversky, 2000; Shafir & LeBoeuf, 2002; Stanovich, 1999, 2004, 2009). These are caused by many well-known cognitive biases: base-rate neglect, framing effects, representativeness biases, anchoring biases, availability bias, outcome bias, and vividness effects, to name just a few."*

Evans & Stanovich 2010

[63] "*In the critical thinking literature, the ability to evaluate evidence and arguments independently of one's prior beliefs and opinions is a skill that is strongly emphasized (Baron, 2008; Dole & Sinatra, 1998; Ennis, 1987, 1996; Kuhn, 2005; Lipman, 1991; Paul, 1984, 1987; Ritchhart & Perkins, 2005; Sternberg, 1997, 2001, 2003; Wade & Tavris, 1993). The disposition toward such unbiased reasoning is almost universally viewed as a characteristic of good thinking. …*

Neimark (1987) emphasizes how associations built up over time will tend to activate a decision for us automatically and unconsciously if we are not reflective and cannot detach from situational cues. The danger of response patterns that are determined too strongly by overlearned cues is a repeated theme in the heuristics and biases literature of cognitive science (e.g., Arkes, 1991; Evans, 2003, 2006, 2007, 2008; Kahneman, 2003; Kahneman & Frederick, 2002; Stanovich, 2003, 2004, 2009; Wilson & Brekke, 1994). … Many tasks in the heuristics and biases branch of the reasoning literature might be said to involve some type of decontextualization skill (Kahneman, 2003; Stanovich, 2003). Tasks are designed to see whether reasoning processes can operate independently of interfering context (world knowledge, prior opinion, vivid examples)."

Evans & Stanovich 2010

[64] "*In the laboratory, the ability to reason in an unbiased manner has been operationalized with a few well-known paradigms. The belief bias syllogism task, one of the most thoroughly explored of these paradigms, shares a number of important features with discourse-logic-based critical thinking measures. This paradigm*

assesses the tendency for judgments of logical validity to be contaminated by prior knowledge of the world—for example, when the validity of a syllogism and the facts expressed in the conclusion of the syllogism conflict (e.g., 'All flowers have petals; roses have petals; therefore, roses are flowers'—which is invalid under necessity). The inability to decouple prior knowledge from reasoning processes has been termed the belief bias effect (Evans, Barston, & Pollard, 1983). It has been the subject of extensive study in the cognitive science literature, and several formal models of how belief bias operates to disrupt syllogistic reasoning have been proposed (De Neys, 2006; Evans & Curtis- Holmes, 2005; Evans & Feeney, 2004; Garnham & Oakhill, 2005; Klauer, Musch, & Naumer, 2000).

Belief bias has also been revealed in paradigms in which participants must evaluate the quality of empirical evidence in a manner not contaminated by their prior opinion on the issue in question. In several studies, Klaczynski (1997) and colleagues (Klaczynski & Gordon, 1996; Klaczynski & Lavallee, 2005; Klaczynski & Robinson, 2000) presented participants with flawed hypothetical experiments that led to conclusions that were either consistent or inconsistent with prior positions and opinions. Participants then critiqued the flaws in the experiments (which were most often badly flawed). Participants found many more flaws when the experiment's conclusions were inconsistent with their prior opinions than when the experiment's conclusions were consistent with their prior opinions and beliefs."

West, Topiak & Stanovich 2008

[65] "*Degrees of rationality can be assessed in terms of the number and severity of such cognitive biases that individuals display. Failure to display a bias becomes a measure of rational thought."*

Stanovic & Stanovich 2010

[66] "*We should emphasize that the large collection of skills and tasks from the heuristics and biases literature that we examined are not represented on traditional critical thinking tests like the WGCTA. For example, the WGCTA and similar genre of critical thinking tests do not assess the use of base rates in probabilistic reasoning; they do not tap knowledge of the law of large numbers or the ability of people to make regressive probabilistic predictions. They do not assess the gambler's fallacy, nor, in general, do they assess covariation detection, denominator neglect, or Bayesian probabilistic updating. In addition, they do not assess a critical aspect of rationality, descriptive invariance, that is the property that causes framing effects when violated (for reviews of the large body of literature on each of these effects, see Gilovich et al., 2002; Nickerson, 2004; Reyna, Lloyd, & Brainerd, 2003; Stanovich, 1999, 2004, 2008b)."*

West, Topiak & Stanovich 2008

[67] "*Virtually all attempts to classify heuristics and biases tasks end up utilizing a dual-process framework because most of the tasks in the literature on heuristics and biases were deliberately designed to pit a heuristically triggered response against a normative response generated by the analytic system. As Kahneman (2000) notes, 'Tversky and I always thought of the heuristics and biases approach as a two-process theory' (p. 682). Since Kahneman and Tversky launched the heuristics and biases approach in the 1970s, a wealth of evidence has accumulated in support of the dual-process approach. Evidence from cognitive neuroscience and cognitive psychology converges on the conclusion that mental functioning can be characterized by two different types of cognition having somewhat different functions and different strengths and weaknesses (Brainerd & Reyna, 2001; Evans, 2003, 2008, 2009; Evans & Over, 1996, 2004; Feldman Barrett, Tugade, & Engle, 2004; Greene, Nystrom, Engell, Darley, & Cohen, 2004; Kahneman & Frederick, 2002, 2005; Lieberman, 2003; McClure, Laibson, Loewenstein, & Cohen, 2004; Metcalfe & Mischel, 1999; Sloman, 1996, 2002; Stanovich, 1999, 2004). There are many such theories (over 20 dual-process theories are presented in a table in Stanovich, 2004) and they have some subtle differences, but they are similar in that all distinguish autonomous from non-autonomous processing. The two types of processing were termed systems in earlier writings, but theorists have been moving toward more atheoretical characterizations; we therefore follow Evans (2009) in using the terms type 1 and type 2 processing. The defining feature of type 1 processing is its autonomy. Type 1 processes are termed autonomous because (a) their execution is rapid, (b) their execution is mandatory when the triggering stimuli are encountered, (c) they do not put a heavy load on central processing capacity (i.e., they do not require conscious attention), (d) they do not depend on input from high-level control systems, and (e) they can operate in parallel without interfering with each other or with type 2 processing. Type 1 processing would include behavioral regulation by the emotions, the encapsulated modules for solving specific adaptive problems that have been posited by evolutionary psychologists, processes of implicit learning, and the automatic firing of overlearned associations (see Evans, 2007, 2008; Stanovich, 2004, 2009). Type 2 processing contrasts with type 1 processing on each of the critical properties that define the latter. Type 2 processing is relatively slow and*

computationally expensive—it is the focus of our awareness. Many type 1 processes can operate at once in parallel, but only one (or a very few) type 2 thoughts can be executing at once—type 2 processing is thus serial processing. Type 2 processing is often language-based. One of the most critical functions of type 2 processing is to override type 1 processing. All of the different kinds of type 1 processing (processes of emotional regulation, Darwinian modules, associative and implicit learning processes) can produce responses that are irrational in a particular context if not overridden. In order to override type 1 processing, type 2 processing must display at least two (possibly related) capabilities. One is the capability of interrupting type 1 processing and suppressing its response tendencies. Type 2 processing thus involves inhibitory mechanisms of the type that have been the focus of recent work on executive functioning (Hasher, Lustig, & Zacks, 2007; Miyake, Fried- man, Emerson, & Witzki, 2000; Zelazo, 2004)."

Stanovic & Stanovich 2010

[68] "*This concern for the efficiency of information processing as opposed to its rationality is mirrored in the status of intelligence tests. They are measures of efficiency but not rationality—a point made clear by considering a distinction that is very old in the field of psychometrics. Psychometricians have long distinguished typical performance situations from optimal (sometimes termed maximal) performance situations (see Ackerman, 1994, 1996; Ackerman & Heggestad, 1997; Ackerman & Kanfer, 2004; see also, Cronbach, 1949; Matthews, Zeidner, & Roberts, 2002). Typical performance situations are unconstrained in that no overt instructions to maximize performance are given and the task interpretation is determined to some extent by the participant. The goals to be pursued in the task are left somewhat open. The issue is what a person would typically do in such a situation, given few constraints. Typical performance measures are measures of the reflective mind—they assess in part goal prioritization and epistemic regulation. In contrast, optimal performance situations are those where the task interpretation is determined externally. The person performing the task is instructed to maximize performance and is told how to do so. Thus, optimal performance measures examine questions of efficiency of goal pursuit—they capture the processing efficiency of the algorithmic mind. All tests of intelligence or cognitive aptitude are optimal performance assessments, whereas measures of critical or rational thinking are often assessed under typical performance conditions. The difference between the algorithmic mind and the reflective mind is captured in another well-established distinction in the measurement of individual differences—the distinction between cognitive ability and thinking dispositions. The former are, as just mentioned, measures of the efficiency of the algorithmic mind. The latter travel under a variety of names in psychology—thinking dispositions or cognitive styles being the two most popular. Many thinking dispositions concern beliefs, belief structure and, importantly, attitudes toward forming and changing beliefs. Other thinking dispositions that have been identified concern a person's goals and goal hierarchy. Examples of some thinking dispositions that have been investigated by psychologists are: actively open-minded thinking, need for cognition (the tendency to think a lot), consideration of future consequences, need for closure, superstitious thinking, and dogmatism (Cacioppo, Petty, & Feinstein 1996; Kruglanski & Webster, 1996; Norris & Ennis, 1989; Schommer-Aikins, 2004; Stanovich, 1999, 2009; Sternberg, 2003; Sternberg & Grigorenko, 1997; Strathman, Gleicher, Boninger, & Scott Edwards, 1994). The literature on these types of thinking dispositions is vast and our purpose is not to review that literature here. It is only necessary to note that the types of cognitive propensities that these thinking disposition measures reflect are the tendency to collect information before making up one's mind, to seek various points of view before coming to a conclusion, to think extensively about a problem before responding, to calibrate the degree of strength of one's opinion to the degree of evidence available, to think about future consequences before taking action, to explicitly weigh pluses and minuses of situations before making a decision, and to seek nuance and avoid absolutism. In short, individual differences in thinking dispositions include assessing variation in people's goal management, epistemic values, and epistemic self-regulation—differences in the operation of reflective mind. They are all psychological characteristics that underpin rational thought and action."*

Stanovic & Stanovich 2010

[69] "*To be rational, a person must have well-calibrated beliefs and must act appropriately on those beliefs to achieve goals—both properties of the reflective mind. The person must, of course, have the algorithmic-level machinery that enables him or her to carry out the actions and to process the environment in a way that allows the correct beliefs to be fixed and the correct actions to be taken. Thus individual differences in rational thought and action can arise because of individual differences in intelligence (the algorithmic mind) or because of individual differences in thinking dispositions (the reflective mind). …*

To think rationally means adopting appropriate goals, taking the appropriate action given one's goals and beliefs, and holding beliefs that are commensurate with available evidence. Standard intelligence tests do not

assess such functions (Perkins, 1995, 2002; Stanovich, 2002, 2009; Sternberg, 2003, 2006). For example, although intelligence tests do assess the ability to focus on an immediate goal in the face of distraction, they do not assess at all whether a person has the tendency to develop goals that are rational in the first place. Likewise, intelligence tests provide good measures of how well a person can hold beliefs in short-term memory and manipulate those beliefs, but they do not assess at all whether a person has the tendency to form beliefs rationally when presented with evidence. And again, similarly, intelligence tests give good measures of how efficiently a person processes information that has been provided, but they do not at all assess whether the person is a critical assessor of information as it is gathered in the natural environment. …

Substantial empirical evidence indicates that individual differences in thinking dispositions and intelligence are far from perfectly correlated. Many different studies involving thousands of subjects (e.g., Ackerman & Heggestad, 1997; Austin & Deary, 2002; Baron, 1982; Bates & Shieles, 2003; Cacioppo et al., 1996; Eysenck, 1994; Goff & Ackerman, 1992; Kanazawa, 2004; Kokis, Macpherson, Toplak, West, & Stanovich, 2002; Zeidner & Matthews, 2000) have indicated that measures of intelligence display only moderate to weak correlations (usually less than .30) with some thinking dispositions (e.g., actively open-minded thinking, need for cognition) and near zero correlations with others (e.g., conscientiousness, curiosity, diligence).

Other important evidence supports the conceptual distinction made here between algorithmic cognitive capacity and thinking dispositions. For example, across a variety of tasks from the heuristics and biases literature, it has consistently found that rational thinking dispositions will predict variance in these tasks after the effects of general intelligence have been controlled (Bruine de Bruin, Parker, & Fischhoff, 2007; Klaczynski, Gordon, & Fauth, 1997; Klaczynski & Lavallee, 2005; Klaczynski & Robinson, 2000; Kokis et al., 2002; Macpherson & Stanovich, 2007; Newstead, Handley, Harley, Wright, & Farrelly, 2004; Parker & Fischhoff, 2005; Sá & Stanovich, 2001; Stanovich & West, 1997, 1998c, 2000; Toplak, Liu, Macpherson, Toneatto, & Stanovich, 2007; Toplak & Stanovich, 2002).

Measures of thinking dispositions tell us about the individual's goals and epistemic values—and they index broad tendencies of pragmatic and epistemic self-regulation at a high level of cognitive control. The empirical studies cited indicate that these different types of cognitive predictors are tapping separable variance, and the reason that this is to be expected is because cognitive capacity measures such as intelligence and thinking dispositions map on to different levels in the tripartite model."

Stanovic & Stanovich 2010

[70] *"A simple example of miserly processing is discussed by Kahneman and Frederick (2002). They describe a simple experiment in which people were asked to consider the following puzzle: 'A bat and a ball cost $1.10 in total. The bat costs $1 more than the ball. How much does the ball cost?'*

Many people offer the response that first comes to mind—10¢– without thinking further and realizing that this cannot be right. The bat would then have to cost $1.10 and the total cost would be $1.20 rather than the required $1.10. People often do not think deeply enough to make this simple correction though, and many students at very selective universities will answer incorrectly and move on to the next problem before realizing that their shallow processing has led them to make an error. Frederick (2005) has found that large numbers of highly selected students at MIT, Princeton, and Harvard, when given this and other similar problems, are cognitive misers like the rest of us. The correlation between intelligence and a set of similar items is quite modest, in the range of .40 to .50 (Gilhooly & Murphy, 2005).

Many other biases of the cognitive miser show correlations no greater than those shown in the Frederick bat-and-ball problem. In fact, some cognitive biases are almost totally dissociated from intelligence. Myside bias, for example, is virtually independent of intelligence (Macpherson & Stanovich, 2007; Sá, Kelley, Ho, & Stanovich, 2005; Stanovich & West, 2007, 2008a, 2008b; Toplak & Stanovich, 2003). Individuals with higher IQs in a university sample are no less likely to process information from an egocentric perspective than are individuals with relatively lower IQs.

Irrational behavior can occur not just because of miserly processing tendencies but also because the right mindware (cognitive rules, strategies, knowledge, and belief systems) is not available to use in decision making."

Stanovic & Stanovich 2010

[71] *"Metacomponents are higher order executive processes used to plan what one is going to do, monitor it while one is doing it, and evaluate it after it is done. The metacomponents include recognising that a problem exists,*

defining the nature of the problem, deciding on a set of steps for solving the problem, ordering these steps into a coherent strategy, deciding upon a form of mental representation for information, allocating one's time and resources in solving a problem, monitoring one's solution to a problem as the problem is being solved, and utilising feedback regarding problem solving after one's problem solving has been completed."

Sternberg 1986

[72] "*Consider, for example, the performance components of induction. These include encoding stimuli, mapping relations between relations, applying relations from one domain to another, justifying potential responses, and responding.*"

Sternberg 1986

[73] "*Three such composants are selective encoding, which involves screening relevant from irrelevant information, selective combination, which involves putting together the relevant information in a coherent and organized way; and selective comparison, which involves relating old, previously known information to new, about to be learned information.*"

Sternberg 1986

[74] "*The lessons are: acquiring expertise in critical thinking is hard; practice in critical thinking skills themselves enhances skills; the transfer of skills must be practiced; some theoretical knowledge is required; diagramming arguments (argument mapping) promotes skills; and students are prone to belief preservation.*"

Gelder 2005

[75] "*Humans are not naturally critical. Indeed, like ballet, critical thinking is a highly contrived activity. Running is natural; nightclub dancing is less so; but ballet is something people can only do well with many years of painful, expensive, dedicated training. Evolution did not intend us to walk on the ends of our toes, and whatever Aristotle might have said, we were not designed to be all that critical either. Evolution does not waste effort making things better than they need to be, and homo sapiens evolved to be just logical enough to survive, while competitors such as Neanderthals and mastodons died out.*"

Gelder 2005

[76] "*From the cognitive scientist's point of view, the mental activities that are typically called critical thinking are actually a subset of three types of thinking: reasoning, making judgments and decisions, and problem solving. I say that critical thinking is a subset of these because we think in these ways all the time, but only sometimes in a critical way. Deciding to read this article, for example, is not critical thinking. But carefully weighing the evidence it presents in order to decide whether or not to believe what it says is. Critical reasoning, decision making, and problem solving—which, for brevity's sake, I will refer to as critical thinking—have three key features: effectiveness, novelty, and self-direction. Critical thinking is effective in that it avoids common pitfalls, such as seeing only one side of an issue, discounting new evidence that disconfirms your ideas, reasoning from passion rather than logic, failing to support statements with evidence, and so on. Critical thinking is novel in that you don't simply remember a solution or situation that is similar enough to guide you. For example, solving a complex but familiar physics problem by applying a multi-step algorithm isn't critical thinking because you are really drawing on memory to solve the problem. But devising a new algorithm is critical thinking. Critical thinking is self-directed in that the thinker must be calling the shots: We wouldn't give a student much credit for critical thinking if the teacher were prompting each step he took.*"

Willingham 2007

[77] "*Can critical thinking actually be taught? Decades of cognitive research point to a disappointing answer: not really. People who have sought to teach critical thinking have assumed that it is a skill, like riding a bicycle, and that, like other skills, once you learn it you can apply it in any situation. Research from cognitive science shows that thinking is not that sort of skill. The processes of thinking are intertwined with the content of thought (that is, domain knowledge).*"

Willingham 2007

[78] "*… if you remind a student to 'look at an issue from multiple perspectives' often enough, he will learn that he ought to do so, but if he doesn't know much about an issue, he can't think about it from multiple perspectives. You can teach students maxims about how they ought to think, but without background knowledge and practice, they probably will not be able to implement the advice they memorize.*"

Willingham 2007

[79] "*Studies of the Philosophy for Children program may be taken as typical. Two researchers identified eight studies that evaluated academic outcomes and met minimal research-design criteria. (Of these eight, only one had been subjected to peer review.) Still, they concluded that three of the eight had identifiable problems that clouded the researchers' conclusions. Among the remaining five studies, three measured reading ability, and one of these reported a significant gain. Three studies measured reasoning ability, and two reported significant gains. And, two studies took more impressionistic measures of student's participation in class (e.g., generating ideas, providing reasons), and both reported a positive effect.*"

Willingham 2007

[80] "*Despite the difficulties and general lack of rigor in evaluation, most researchers reviewing the literature conclude that some critical thinking programs do have some positive effect. But these reviewers offer two important caveats. First, as with almost any educational endeavor, the success of the program depends on the skill of the teacher. Second, thinking programs look good when the outcome measure is quite similar to the material in the program. As one tests for transfer to more and more dissimilar material, the apparent effectiveness of the program rapidly drop*s."

Willingham 2007

[81] "*If knowledge of how to solve a problem never transferred to problems with new surface structures, schooling would be inefficient or even futile—but of course, such transfer does occur. When and why is complex, but two factors are especially relevant for educators: familiarity with a problem's deep structure and the knowledge that one should look for a deep structure.*"

W Willingham 2007

"Here's an example: A treasure hunter is going to explore a cave up on a hill near a beach. He suspected there might be many paths inside the cave so he was afraid he might get lost. Obviously, he did not have a map of the cave; all he had with him were some common items such as a flashlight and a bag. What could he do to make sure he did not get lost trying to get back out of the cave later? The solution is to carry some sand with you in the bag, and leave a trail as you go, so you can trace your path back when you're ready to leave the cave. About 75 percent of American college students thought of this solution—but only 25 percent of Chinese students solved it.6 The experimenters suggested that Americans solved it because most grew up hearing the story of Hansel and Gretel, which includes the idea of leaving a trail as you travel to an unknown place in order to find your way back. The experimenters also gave subjects another puzzle based on a common Chinese folk tale, and the percentage of solvers from each culture reversed. (To read the puzzle based on the Chinese folk tale, and the tale itself, go to www.aft.org/ pubs-reports/american_educator/index.htm.) It takes a good deal of practice with a problem type before students know it well enough to immediately recognize its deep structure, irrespective of the surface structure, as Americans did for the Hansel and Gretel problem."

Willingham 2007

"They are little chunks of knowledge—like 'look for a problem's deep structure' or 'consider both sides of an issue'—that students can learn and then use to steer their thoughts in more productive directions." « Le problème de ce genre de stratégie métacognitive est qu'elle ne peut pas porter beaucoup plus loin qu'à amener à se répéter au bon moment la stratégie métacognitive. » "Thus, a student who has been encouraged many times to see both sides of an issue, for example, is probably more likely to spontaneously think 'I should look at both sides of this issue' when working on a problem. … Unfortunately, metacognitive strategies can only take you so far. Although they suggest what you ought to do, they don't provide the knowledge necessary to implement the strategy."

Willingham 2007

"Understanding and using conditional probabilities is essential to scientific thinking because it is so important in reasoning about what causes what. But people's success in thinking this way depends on the particulars of how the question is presented. Studies show that adults sometimes use conditional probabilities successfully, but fail to do so with many problems that call for it. Even trained scientists are open to pitfalls in reasoning about conditional probabilities (as well as other types of reasoning). Physicians are known to discount or misinterpret new patient data that conflict with a diagnosis they have in mind, and Ph.D.-level scientists are prey to faulty reasoning when faced with a problem embedded in an unfamiliar context."

Willingham 2007

[82] "*But critical thinking is very different. As we saw in the discussion of conditional probabilities, people can engage in some types of critical thinking without training, but even with extensive training, they will sometimes fail to think critically. This understanding that critical thinking is not a skill is vital.*"

Willingham 2007

[83] "*It tells us that teaching students to think critically probably lies in small part in showing them new ways of thinking, and in large part in enabling them to deploy the right type of thinking at the right time...*

What do all these studies boil down to? First, critical thinking (as well as scientific thinking and other domain-based thinking) is not a skill. There is not a set of critical thinking skills that can be acquired and deployed regardless of context. Second, there are metacognitive strategies that, once learned, make critical thinking more likely. Third, the ability to think critically (to actually do what the metacognitive strategies call for) depends on domain knowledge and practice."

Willingham, D. T. (2007). Critical thinking: Why it is so hard to teach? *American federation of teachers summer 2007, p. 8-19.*

[84] "*For example, in one experiment researchers taught college students principles for evaluating evidence in psychology studies—principles like the difference between correlational research and true experiments, and the difference between anecdote and formal research (Bensley & Spero, 2014). These principles were incorporated into regular instruction in a psychology class, and their application was practiced in that context. Compared to a control group that learned principles of memory, students who learned the critical thinking principles performed better on a test that required evaluation of psychology evidence.*

There is even evidence that critical thinking skills can be taught and applied in complex situations under time pressure. In one experiment, officers in the Royal Netherlands Navy received training in the analysis of complex battlefield problems in a high- fidelity tactical simulator. They were first taught a sequence of steps to undertake when analyzing this sort of problem, and then underwent a total of 8 hours of training on surface warfare problems, with feedback from an expert. The critical outcome measure was performance (without feedback) in a new surface warfare problem, as well as performance on air warfare problems. Judges assessed the quality of participant's action contingency plans, and those receiving the training outperformed control subjects (Helsdingen et al., 2010).

When we think of critical thinking, we think of something bigger than its domain of training. When I teach students how to evaluate the argument in a set of newspaper editorials, I am hoping that they will learn to evaluate arguments generally, not just those they read, and not just those they would find in other editorials. This aspect of critical thinking is called transfer, and the research literature evaluating how well critical thinking skills transfer to new problems is decidedly mixed. …

It is not useful to think of critical thinking skills, once acquired, as broadly applicable. Wanting students to be able to 'analyse, synthesise and evaluate' information sounds like a reasonable goal. But analysis, synthesis, and evaluation mean different things in different disciplines. Literary criticism has its own internal logic, its norms for what constitutes good evidence and a valid argument. These norms differ from those found in mathematics, for example. And indeed, different domains—science and history, say—have different definitions of what it means to 'know' something. Thus, our goals for student critical thinking must be domain-specific. An overarching principle like 'think logically' is not a useful goal. …

Experimental evidence shows that an expert does not think as well outside her area of expertise, even in a closely related domain. She is still better than a novice, but her skills do not transfer completely. For example, knowledge of medicine transfers poorly among subspecialties; neurologists do not diagnose cardiac cases well (Rikers, Schmidt, & Boshuizen, 2002). Expertise in writing is similarly encapsulated; a technical writer who specialises in writing instruction pamphlets for home electronics can't write newspaper articles (Kellogg, 2018). Perhaps most surprisingly the analytic abilities of professional philosophers do not extend to everyday judgments. Philosophers are no less susceptible than average adults to being swayed by irrelevant features of problems like question order or wording (Schwitzgebel & Cushman, 2015)."

Willingham 2019

[85] "*First, identify what is meant by critical thinking in each domain. Be specific. What tasks showing critical thinking should a high school graduate be able to do in mathematics, history, and other subjects? It is not useful*

to set a goal that students 'think like historians,' or 'learn the controversies surrounding historical events.' If students are to read as historians do, they need to learn specific skills like interpreting documents in light of their sources, corroborating them, and putting them in historical context. Notably, skillful reading is different in other disciplines. Scientists believe that the source of a document is irrelevant so long as it is trustworthy. And unlike historical documents, scientific documents are written in a consistent format. Learning to read like a scientist means, in part, learning the conventions of this format. These skills should be explicitly taught and practiced—there is evidence that simple exposure to this sort of work without explicit instruction is less effective (Abrami et al., 2008; Halpern, 1998; Heijltjes, Van Gog, & Paas, 2014). In addition, it is clear that educators will have to pick and choose which skills their students will learn. Even across the long thirteen years of schooling, time is limited. Second, identify the domain content that students must know. We have seen that domain knowledge is a crucial driver of thinking skills. For example, sourcing historical documents means interpreting their content in light of the author, the intended audience, and circumstances under which the author wrote. It is not enough to know that a letter was written by an army sergeant to his wife just before the Battle of Romani. The student must know enough about the historical context to understand how this sourcing information ought to influence his or her interpretation of the letter. Fourth, educators must decide which skills should be revisited across years. Studies show that even if content is learned quite well over the course of half of a school year, about half will be forgotten in three years (Pawl et al., 2012). That doesn't mean there's no value in exposing students to content just once; most students will forget much but they'll remember something, and for some students, an interest may be kindled. But when considering skills we hope will stick with students for the long term, we should plan on at least three to five years of practice (Bahrick, 1984; Bahrick & Hall, 1991). Most of the time, this practice will look different—it will be embedded in new skills and content. But this revisiting should be assured and planned."

Willingham 2019

[86] *Another area of disagreement among critical thinking researchers is the extent to which critical thinking skills and abilities can be transferred to new contexts. For example, researchers have noted that students may exhibit critical thinking skills and abilities in one context, or domain, but fail to do so in another (Willingham, 2007). This issue is closely related to that of the domain-specificity of critical thinking. For example, those maintaining that critical thinking is completely domain-specific are more likely to be skeptical of students' abilities to transfer critical thinking skills from one domain to another (Ennis, 1989). Accepted wisdom within cognitive psychology holds that spontaneous transfer to new contexts is rare (Kennedy et al., 1991; Pithers & Soden, 2000; Willingham, 2007). Others, however, are more sanguine about the possibility of student transfer, particularly if students are given opportunities to practice critical thinking skills in multiple domains and contexts and if students are taught specifically to transfer those skills (Kennedy et al., 1991). McPeck (1990), a staunch proponent of domain specificity, notes that his approach does not preclude the transfer of critical thinking skills and abilities to real-world contexts, particularly when instruction emphasizes authentic learning activities that represent problems encountered in daily life.*

Lai 2011

[87] *Existing published assessments of critical thinking are numerous, and include the California Critical Thinking Skills Test (Facione, 1990), the Cornell Critical Thinking Tests (Ennis & Millman, 2005), the Ennis-Weir Critical Thinking Essay Test (Ennis & Weir, 1985), and the Watson-Glaser Critical Thinking Appraisal (Watson & Glaser, 1980). As Ku (2009) points out, these instruments vary widely in both purpose and item format. However, as Kennedy et al. (1991) note, none of these tests are intended for use with students below the fourth-grade level. Moreover, these assessments tend to be general critical thinking assessments rather than subject-specific.*

Lai 2011

[88] *Despite differences among the three schools of thought and their approaches to defining critical thinking, there exist areas for agreement. First, researchers of critical thinking typically agree on the specific abilities encompassed by the definition, which include:*

- *analyzing arguments, claims, or evidence (Ennis, 1985; Facione, 1990; Halpern, 1998; Paul, 1992);*
- *making inferences using inductive or deductive reasoning (Ennis, 1985; Facione, 1990; Paul, 1992; Willingham, 2007);*
- *judging or evaluating (Case, 2005; Ennis, 1985; Facione, 1990; Lipman, 1988; Tindal & Nolet, 1995); and*
- *making decisions or solving problems (Ennis, 1985; Halpern, 1998; Willingham, 2007).*

Other abilities or behaviors identified as relevant to critical thinking include asking and answering questions for clarification (Ennis, 1985); defining terms (Ennis, 1985); identifying assumptions (Ennis, 1985; Paul, 1992); interpreting and explaining (Facione, 1990); reasoning verbally, especially in relation to concepts of likelihood and uncertainty (Halpern, 1998); predicting (Tindal & Nolet, 1995); and seeing both sides of an issue (Willingham, 2007).

Lai 2011

[89] *Some researchers have argued that the link between critical thinking and metacognition is self-regulation. For example, the APA Delphi report includes self-regulation as one component skill of critical thinking (Facione, 1990). Schraw et al. (2006) draw connections between metacognition, critical thinking, and motivation under the umbrella of self-regulated learning, which they define as "our ability to understand and control our learning environments" (p. 111). Self-regulated learning, in turn, is seen as comprising three components: cognition, metacognition, and motivation. The cognitive component includes critical thinking, which Schraw and associates explain consists of identifying and analyzing sources and drawing conclusions. However, others have argued that critical thinking and metacognition are distinct constructs. For example, Lipman (1988) has pointed out that metacognition is not necessarily critical, because one can think about one's thought in an unreflective manner. McPeck, on the other hand, argues that the ability to recognize when a particular skill is relevant and to deploy that skill is not properly a part of critical thinking but actually represents general intelligence (1990). At the very least, metacognition can be seen as a supporting condition for critical thinking, in that monitoring the quality of one's thought makes it more likely that one will engage in high-quality thinking.*

Lai 2011

[90] *Many researchers working in the area of critical thinking lament the poor state of critical thinking in most educated adults and children. For example, Halpern (1998) points to research from the field of psychology, concluding that many, if not most, adults fail to think critically in many situations. Kennedy et al., (1991) and Van Gelder (2005) have likewise concluded that many adults lack basic reasoning skills. Halpern (1998) cites the example that large numbers of people profess to believe in paranormal phenomena, despite a lack of evidence in support of such things. Halpern attributes such failures not to the inability to reason well but to simple "bugs" in reasoning. She argues that human beings are programmed to look for patterns, particularly in the form of cause-and-effect relationships, even when none exist. Van Gelder (2005) echoes this sentiment, characterizing humans as "pattern-seekers and story-tellers" (p. 42). This inclination results in a tendency to jump to the first explanation that makes intuitive sense without carefully scrutinizing alternative possibilities, a phenomenon that Perkins, Allen, & Hafner (1983) have termed "makes-sense epistemology" (p. 286). Moreover, the general public often finds "personal experience" to be more compelling evidence than a carefully conducted, scientific study. Given these natural tendencies toward deficient reasoning, Halpern warns that we should not expect to see dramatic improvements in critical thinking over time as a result of instructional interventions. Improvements in critical thinking, when they do occur, are slow and incremental (Halpern, 1998).*

Lai 2011

[91] *According to Kuhn's (1999) theoretical framework, metacognitive knowing characterizes the first stirrings of critical thought in very young children. There are two distinct stages within metacognitive knowing. The first stage is called Realism and is typically achieved between the ages of 3 and 5. This stage is characterized by the belief that assertions are expressions of someone's belief, and as such, may depart from reality. Thus, the child is able to identify true and false statements. Prior to reaching this stage, children regard beliefs and assertions as isomorphic with reality. "In other words, the world is a simple one in which things happen and we can tell about them. There are no inaccurate renderings of events" (p. 19).*

According to Kuhn's framework (1999), the second stage of metacognitive knowing, typically achieved by 6 years of age, allows the child to be aware of sources of knowledge and further, to distinguish between theory and evidence. In other words, prior to reaching this second stage, the child has difficulty distinguishing evidence for the claim that an event has occurred from the causal theory that makes occurrence of the event plausible. In other words, is something true because it makes intuitive sense or because there is empirical evidence for it? Kuhn describes a study (Kuhn & Pearsall, 1998) in which children were shown a series of pictures depicting two runners competing in a race. The last picture shows one of the runners holding up a trophy and smiling. When children were asked who won the race, most children correctly indicated that the runner represented in the final photo was the winner. However, when asked to justify this claim, younger children tended to cite causal theories ("because he is wearing fast shoes") rather than evidence in support of the claim ("because he is holding a

trophy"). According to Kuhn, by the second stage of metacognitive knowing children are able to make this distinction.

Based on the empirical research in meta-memory, Kuhn's framework (1999) also portrays meta-strategic knowing in two stages. According to Kuhn, during the first stage, typically achieved during middle childhood, children begin to understand the value of cognitive strategies in aiding cognition. A child who has reached this stage recognizes that a memory strategy such as categorization will aid recall and tends to effectively manage and deploy cognitive resources during problem solving (Kuhn, 1999). The second stage of meta-strategic knowing may not be achieved at all. If it is attained, it is typically reached during adolescence and adulthood. According to Kuhn, this stage is characterized by consistent and appropriate strategy selection from a repertoire of available strategies. Thus, the individual monitors strategy.

Based on the empirical research in meta-memory, Kuhn's framework (1999) also portrays meta-strategic knowing in two stages. According to Kuhn, during the first stage, typically achieved during middle childhood, children begin to understand the value of cognitive strategies in aiding cognition. A child who has reached this stage recognizes that a memory strategy such as categorization will aid recall and tends to effectively manage and deploy cognitive resources during problem solving (Kuhn, 1999). The second stage of meta-strategic knowing may not be achieved at all. If it is attained, it is typically reached during adolescence and adulthood. According to Kuhn, this stage is characterized by consistent and appropriate strategy selection from a repertoire of available strategies. Thus, the individual monitors strategy According to Kuhn (1999), the second stage in epistemological understanding, labeled the Multiplist Epistemological position, tends to be prevalent during adolescence. During this stage, the individual acknowledges that experts can disagree and actually relinquishes the idea of certainty. A person in this stage moves to the opposite end of the subjectivity-objectivity continuum, vis-à-vis those in the Absolutist stance. Instead of viewing the world as inherently and objectively knowable, individuals in this stage perceive the world as a completely subjective place. In other words, "because all people have a right to their opinions, all opinions are equally right" (p. 22). Kuhn points out that many people become permanently stuck in this phase.

Finally, Kuhn (1999) argues that the last stage in epistemological understanding (and critical thinking), to which only a minority of people will ever progress, is known as Epistemological Metaknowing. According to Kuhn's framework (1999), at this stage the individual is able to balance the subjective and objective, recognizing a multiplicity of valid This person uses judgment, evaluation, and argumentation to sift through opinions and arrive at those that are most valid. Not all opinions are valued equally; rather, reason, logic, and empirical evidence can be used to privilege certain positions over others (Kuhn, 1999). This person uses judgment, evaluation, and argumentation to sift through opinions and arrive at those that are most valid. Not all opinions are valued equally; rather, reason, logic, and empirical evidence can be used to privilege certain positions over others (Kuhn, 1999).

Lai 2011

[92] *Reasoning is generally seen as a means to improve knowledge and make better decisions. However, much evidence shows that reasoning often leads to epistemic distortions and poor decisions. This suggests that the function of reasoning should be rethought. Our hypothesis is that the function of reasoning is argumentative. It is to devise and evaluate arguments intended to persuade. Reasoning so conceived is adaptive given the exceptional dependence of humans on communication and their vulnerability to misinformation. A wide range of evidence in the psychology of reasoning and decision making can be reinterpreted and better explained in the light of this hypothesis. Poor performance in standard reasoning tasks is explained by the lack of argumentative context. When the same problems are placed in a proper argumentative setting, people turn out to be skilled arguers. Skilled arguers, however, are not after the truth but after arguments supporting their views. This explains the notorious confirmation bias. This bias is apparent not only when people are actually arguing, but also when they are reasoning proactively from the perspective of having to defend their opinions. Reasoning so motivated can distort evaluations and attitudes and allow erroneous beliefs to persist. Proactively used reasoning also favors decisions that are easy to justify but not necessarily better. In all these instances traditionally described as failures or flaws, reasoning does exactly what can be expected of an argumentative device: Look for arguments that support a given conclusion, and, ceteris paribus, favor conclusions for which arguments can be found.*

Mercier & Sperber 2011

[93] *We understand critical thinking to be purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological,*

criteriological, or contextual considerations upon which that judgment is based. CT is essential as a tool of inquiry. As such, CT is a liberating force in education and a powerful resource in one's personal and civic life. While not synonymous with good thinking, CT is a pervasive and self-rectifying human phenomenon. The ***ideal critical thinker i****s habitually inquisitive, well-informed, trustful of reason, open-minded, flexible, fair- minded in evaluation, honest in facing personal biases, prudent in making judgments, willing to reconsider, clear about issues, orderly in complex matters, diligent in seeking relevant information, reasonable in the selection of criteria, focused in inquiry, and persistent in seeking results which are as precise as the subject and the circumstances of inquiry permit. Thus, educating good critical thinkers means working toward this ideal. It combines developing CT skills with nurturing those dispositions which consistently yield useful insights and which are the basis of a rational and democratic society.*

Facione, 1990

[94] "*Communication brings vital benefits, but carries a major risk for the audience of being accidentally or intentionally misinformed. Nor is there any failsafe way of calibrating one's trust in communicated information so as to weed out all and only the misinformation. Given that the stakes are so high, it is plausible that there has been ongoing selective pressure in favour of any available cost-effective means to least approximate such sorting. Since there are a variety of considerations relevant to the granting or withholding of epistemic trust, we will explore the possibility that different abilities for epistemic vigilance may have emerged in biological and cultural evolution, each specialising in a particular kind of relevant considerations. Factors affecting the acceptance or rejection of a piece of communicated information may have to do either with the source of the information—who to believe; or with its content—what to believe.*"

Sperber et al 2010

[95] *"We claim that humans have a suite of cognitive mechanisms for epistemic vigilance, targeted at the risk of being misinformed by others. Here we present this claim and consider some of the ways in which epistemic vigilance works in mental and social life. Our aim is to integrate into a coherent topic for further research a wide range of assumptions developed elsewhere by ourselves or others, rather than to present detailed arguments for each." ...*

How reliable are others as sources of information? In general, they are mistaken no more often than we are—after all, 'we' and 'they' refer to the same people—and they know things that we don't know. So it should be advantageous to rely even blindly on the competence of others. Would it be more advantageous to modulate our trust by exercising some degree of vigilance towards the competence of others? That would depend on the cost and reliability of such vigilance. But in any case, the major problem posed by communicated information has to do not with the competence of others, but with their interests and their honesty. While the interests of others often overlap with our own, they rarely coincide with ours exactly. In a variety of situations, their interests are best served by misleading or deceiving us. It is because of the risk of deception that epistemic vigilance may be not merely advantageous but indispensable if communication itself is to remain advantageous. ...

People stand to gain immensely from communication with others, but this leaves them open to the risk of being accidentally or intentionally misinformed, which may reduce, cancel, or even reverse these gains. The fact that communication is so pervasive despite this risk suggests that people are able to calibrate their trust well enough to make it advantageous on average to both communicator and audience (Sperber, 2001; Bergstrom et al., 2006). For this to happen, the abilities for overt intentional communication and epistemic vigilance must have evolved together, and must also develop together and be put to use together.

Sperber et al 2010

[96] *"Some contents are intrinsically believable even if they come from an untrustworthy source. Examples include tautologies, logical proofs, truisms, and contents whose truth is sufficiently evidenced by the act of communication itself (e.g. saying, 'Je suis capable de dire quelques mots en français'). Other contents are intrinsically unbelievable even if they come from a trustworthy source. Examples include logical contradictions, blatant falsehoods, and contents whose falsity is sufficiently evidenced by the act of communication itself (e.g. saying, 'I am mute'). ... checking takes place against the narrow context of beliefs used in the search for a relevant interpretation of the utterance."*

Sperber et al 2010

[97] *"Judgements about the trustworthiness of informants may be more or less general or contextualised. You may think, 'Mary is a trustworthy person', meaning it both epistemically and morally, and therefore expecting what*

Mary says to be true, what she does to be good, and so on. Or you may trust (or mistrust) someone on a particular topic in specific circumstances: 'You can generally trust Joan on Japanese prints, but less so when she is selling one herself'. Trust can be allocated in both these ways, but how do they compare from a normative point of view ?

A reliable informant must meet two conditions: she must be competent, and she must be benevolent. That is, she must possess genuine information (as opposed to misinformation or no information), and she must intend to share that genuine information with her audience (as opposed to making assertions she does not regard as true, through either indifference or malevolence). Clearly, the same informant may be competent on one topic but not on others, and benevolent towards one audience in certain circumstances, but not to another audience or in other circumstances. This suggests that trust should be allocated to informants depending on the topic, the audience, and the circumstances. However such precise calibration of trust is costly in cognitive terms, and, while people are often willing to pay the price, they also commonly rely on less costly general impressions of competence, benevolence and overall trustworthiness."

Sperber et al 2010

[98] *There is a growing body of research on the development of children's epistemic vigilance (for reviews, see e.g. Koenig and Harris, 2007; Heyman, 2008; Clement, in press; Corriveau and Harris, in press; Nurmsoo et al., in press). This shows that even at a very early age, children do not treat all communicated information as equally reliable. At 16 months, they notice when a familiar word is inappropriately used (Koenig and Echols, 2003). By the age of two, they often attempt to contradict and correct assertions that they believe to be false (e.g. Pea, 1982). These studies challenge the widespread assumption that young children are simply gullible*

Sperber et al 2010

[99] *"One group of 3-, 4- and 5-year olds watched and listened as two speakers narrated a short passage from the story of 'Curious George'. One spoke English with a native (North-American) accent. The other spoke English with a foreign (Spanish) accent. A second group of children of the same age watched and listened as the two speakers narrated a short passage from 'Jabberwocky'—the nonsense poem by Lewis Carroll. Although syntactically well-formed, the sentences in this passage were not meaningful so that any differences in trust following this induction could not be attributed to differential comprehension of the two speakers. Following both types of induction, children were given an opportunity to seek and endorse information about the use of four unfamiliar artefacts from the two speakers. They offered conflicting demonstrations of how to use any given artefact. For example, one speaker looked through a plastic sprinkler attachment as if it were a telescope, whereas the other speaker held it to her mouth and blew in it. Children preferred to seek and endorse information from the native-accented speaker. This preference was equally strong in all three age groups and equally strong following the meaningful, 'Curious George' induction and the meaningless, 'Jaberwocky' induction. Note that the induction phase and the test phase of this experiment differed in both modality and domain. The induction phase involved audible differences in accent. The test phase involved visible differences in tool use. Nevertheless, children used the audible cues of group membership to guide their learning about tool use."*

Harris & Corriveau 2011

[100] *"...suppose that children encounter two informants who make conflicting claims that are novel and therefore impossible for children to adjudicate themselves. However, the claims made by one informant elicit approval from bystanders, whereas the claims made by the other elicit disapproval. Do children use such bystanders' reactions to moderate their trust in the novel claims made by each informant? To examine this possibility, we had 4-year olds watch as two speakers produced conflicting names for a series of unfamiliar objects. For example, faced with the sprinkler attachment, one speaker might call it a 'feppin' and the other might call it a 'merval'. The two bystanders reacted differently to the two speakers. Having listened to one, they nodded and smiled. Having listened to the other, they shook their head and frowned. Subsequently, children were asked for their judgement. They were reminded that one speaker had called it a feppin and the other had called it a merval—what did they think? Children overwhelmingly endorsed the speaker who had attracted bystanders' approval rather than disapproval. In the next stage of the experiment, we tested if children would continue to regard the speaker who had received bystanders' approval as more trustworthy even in the absence of any feedback from the bystanders. To assess this possibility, the two bystanders left the room, and testing continued as before with the two informants making conflicting claims about unfamiliar objects. Children continued to display selective trust in the two speakers—they were more likely to endorse the names supplied by the speaker who had received bystanders' approval even though, at this point in the experiment, the bystanders were no*

longer present and could supply no cues. By implication, the cultural typicality of the two speakers—the extent to which their claims had met with approval versus disapproval—led children to regard one of them as a more trustworthy informant. However, an alternative interpretation of these results is that children did not conclude that the two informants differed in terms of cultural typicality but in terms of likeability. After all, in expressing their approval, the bystanders had smiled at one informant, and in expressing their disapproval, they had frowned at the other. Arguably, children preferred to endorse the speaker whom they inferred to be more likeable, as indexed by the bystanders' reactions. In a follow-up study, we again had two informants as well as an additional pair of adults who sided with one informant and not with the other. However, we altered the way in which this endorsement was expressed. Several unfamiliar objects were set out on a table and the experimenter asked the adults to say which of them was, for example, 'a slod'. Three of the adults pointed to the same object, whereas the fourth—the lone dissenter—pointed to a different object. This pattern was repeated for four trials with the same person always in the role of a lone dissenter. After watching the adults' responses, children were invited to express their view. As in the previous study, children strongly favoured the majority view, effectively endorsed by three of the adults, as opposed to the minority view endorsed by only a single adult.

Harris & Corriveau 2011

[101] *"No matter how non-selective children are in what they learn from others, they are selective in whom they learn from. We have identified two broad classes of heuristics—one class helps children to select among informants with whom they have had previous interactions, and the second class helps children to differentiate among relatively unfamiliar informants whom they have just met...*

Within the first class, children display two biases. First, they display a preference for the information supplied by a familiar caregiver versus a stranger (provided that they have not developed an avoidant relationship with that caregiver). Second, children prefer information supplied by someone who has proven to be a reliable source of information in the past. Taken together, these two biases are likely to converge on a proclivity for vertical cultural learning—a bias to endorse and imitate the claims and demonstrations of adults who have a record of providing reliable care, accurate information, or both. ... The second class of biases enables children to differentiate among informants with whom they have had no protracted interaction. As noted, this class leads children to prefer informants who appear to be culturally typical, either in the sense that the informants signal that they belong to the same group as the children (because of the way that they speak or look) or in the sense that other potential informants assent to, rather than dissent from, the information offered by the informant. These biases are likely to promote oblique and horizontal cultural learning that is relatively conservative. When children encounter someone who is not a familiar caregiver, they will be more inclined to accept guidance from that person if he or she appears to belong to, and receives endorsement from, the children's cultural group. Stated differently, children's receptivity to both oblique and horizontal learning does not extend to all-comers. They are less likely to trust information that is provided by members of another cultural group or by deviants from within their own group."

Harris & Corriveau 2011

[102] *Epistemic vigilance directed at informants yields a variety of epistemic attitudes (acceptance, doubt or rejection, for instance) to the contents communicated by these informants. There is some evidence that three-year-old children are aware of attitudes such as endorsement or doubt (Fusaro and Harris, 2008), and are also aware that assertions can be stronger or weaker (Sabbagh and Baldwin, 2001; Birch et al., 2008; Matsui et al., 2009). Children are able to make sense of comments on the reliability of what is communicated (e.g. Fusaro and Harris, 2008, Clement et al., 2004). As a result, they can take advantage of the epistemic judgments of others, and enrich their own epistemological understanding and capacity for epistemic vigilance in doing so.*

Sperber et al 2010

[103] *"Likewise, the acquisition of other metacognitive and metastrategic skills is a gradual process. Early strategies for coordinating theory and evidence are replaced with better ones, but there is not a stage-like change from using an older strategy to a newer one. … However, metastrategic competence does not appear to routinely develop in the absence of instructions*

Morris et al. 2012

[104] *"A reliable informant must meet two conditions: she must be competent, and she must be benevolent. That is, she must possess genuine information (as opposed to misinformation or no information), and she must intend to share that genuine information with her audience (as opposed to making assertions she does not regard as true, through either indifference or malevolence). Clearly, the same informant may be competent on one topic but not*

on others, and benevolent towards one audience in certain circumstances, but not to another audience or in other circumstances. This suggests that trust should be allocated to informants depending on the topic, the audience, and the circumstances. However such precise calibration of trust is costly in cognitive terms, and, while people are often willing to pay the price, they also commonly rely on less costly general impressions of competence, benevolence and overall trustworthiness."

Sperber et al 2010

[105] *A striking illustration of the tendency to form general judgments of trustworthiness on the basis of very limited evidence is provided in a study by Willis and Todorov (2006). Participants were shown pictures of faces, for either a mere 100 milliseconds or with no time limit, and asked to evaluate the person's trustworthiness, competence, likeability, aggressiveness and attractiveness. Contrary to the authors' expectations, the correlation between judgments with and without time limit was not greater for attractiveness —which is, after all, a property of a person's appearance—than for trustworthiness, while the correlations for aggressiveness and competence were a relatively low. One might wonder if such split-second judgments of trustworthiness have any basis at all, but what this experiment strongly suggests is that looking for signs of trustworthiness is one of the first things we do when we see a new face (see also Ybarra et al., 2001). There is a considerable social psychology literature suggesting that people's behaviour is determined to a significant extent not by their character but by the situation (Ross and Nisbett, 1991; Gilbert and Malone, 1995). If so, judging that someone is generally trustworthy may be a case of the 'fundamental attribution error' (Ross, 1977): that is, the tendency, in explaining or predicting someone's behaviour, to overestimate the role of psychological dispositions and underestimate situational factors. But even without appealing to character psychology, it is possible to defend the view that some people are more generally trustworthy than others, and are to some extent recognisable as such."*

Sperber et al 2010

[106] *"The very social success which is almost a defining feature of cultural information might suggest that (except in cases of cultural conflict) it is uncritically accepted. We will argue, however, that here too epistemic vigilance is at work, but that it needs appropriate cultural and institutional development to meet some of the epistemic challenges presented by cultural information... Often, information spreads through a group from a single source, and is accepted by people along the chains of transmission because they trust the source rather than because of any evidence or arguments for the content. If so, the crucial consideration should be the trustworthiness of the original source. If each person who passes on the information has good independent reasons for trusting the source, this should give people further along the chain good reasons for also trusting the source, and thus for accepting the content originally conveyed. However, people's reasons for trusting the source are in general no more independent of one another than their reasons for accepting the content...*

It might seem, then, that people are simply willing, or even eager, to accept culturally transmitted information without exercising ordinary epistemic vigilance towards it. Boyd, Richerson and Henrich have argued that there is an evolved conformist bias in favour of adopting the behaviour and attitudes of the majority of members of one's community (e.g. Boyd and Richerson, 1985; Henrich and Boyd, 1998). Csibra and Gergely (2009) have argued that people in general, and children in particular, are eager to acquire cultural information, and that this may bias them towards interpreting (and even over-interpreting) communicated information as having cultural relevance, and also towards accepting it. An alternative (or perhaps complementary) hypothesis is that people do exercise some degree of epistemic vigilance towards all communicated information, whether local or cultural, but that their vigilance is directed primarily at information originating in face to face interaction, and not at information propagated on a larger scale. For instance, people may be disposed to pay attention to the problems raised by the non-independence of testimonies, or by discrepancies in their contents, when they are blatantly obvious, as they often are when they occur in face to face interaction, but not otherwise. On a population scale, these problems can remain unnoticed although, on reflection, they are likely to be pervasive. All kinds of beliefs widely shared in the community may propagate throughout a culture by appealing to individual trust in converging testimonies. The trust is not blind, but the epistemic vigilance which should buttress it is short-sighted...

So far, the picture we have sketched of epistemic vigilance on a population scale is somewhat grim. Mechanisms for epistemic vigilance are not geared to filtering information transmitted on such a large scale. Even if we are right to claim that these mechanisms exist, they do not prevent mistaken ideas, undeserved reputations and empty creeds from invading whole populations. However, we did note that it is important not to jump from the fact that people are seriously, even passionately, committed to certain ideas, and expect others to be similarly

committed, to the conclusion that the commitment involved is clearly epistemic. It may be that the content of the ideas matters less to you than who you share them with, since they may help define group identities. When what matters is the sharing, it may be that contents which are unproblematically open to epistemic evaluation would raise objections within the relevant social group, or would be too easily shared beyond that group. So, semi-propositional contents which can be unproblematically accepted by just the relevant group may have a cultural success which is negatively correlated with their epistemic value."

Sperber et al 2010

[107] *"In a number of domains, there are institutional procedures for evaluating the competence of individuals, making these evaluations public through some form of certification, and sanctioning false claims to being so certified. Medical doctors, professors, judges, surveyors, accountants, priests, and so on are generally believed to be experts in their field because they have shown strong evidence of their expertise to experts who are even more qualified. Of course, these procedures may be inadequate or corrupt, and the domain may itself be riddled with errors; but still, such procedures provide clear and easily accessible evidence of an individual's expertise... The institutional organisation of epistemic vigilance is nowhere more obvious than in the sciences, where observational or theoretical claims are critically assessed via social processes such as laboratory discussion, workshops, conferences, and peer review in journals. The reliability of a journal is itself assessed through rankings, and so on (Goldman, 1999). Social mechanisms for vigilance towards the source and vigilance towards the content interact in many ways. In judicial proceedings, for instance, the reputation of the witness is scrutinised in order to strengthen or weaken her testimony. In the sciences, peer review is meant to be purely content-oriented, but is influenced all too often by the authors' prior reputation (although blind reviewing is supposed to suppress this influence), and the outcome of the reviewing process in turn affects the authors' reputation. Certification of expertise, as in the granting of a PhD, generally involves multiple complex assessments from teachers and examiners, who engage in discussion with the candidate and among themselves; these assessments are compiled by educational institutions which eventually deliver a reputation label, 'PhD', for public consumption."*

Sperber et al 2010

[108] *"Our main aim in doing so is to suggest that, to a significant extent, these social mechanisms are articulations of psychological mechanisms linked through extended chains of communication, and, in some cases through institutional patterning (Sperber, 1996). In these population scale articulations, psychological mechanisms combine with cognitive artefacts (e.g. measuring instruments), techniques (e.g. statistical tests of confidence), and procedures (e.g. for cross-examination) to yield distributed epistemic assessment systems (Heintz, 2006) which should be seen as a special kind of distributed cognitive system (Hutchins, 1996)."*

Sperber et al 2010

[109] *"The form of gullibility that is the main target here has the three following traits. First, it views gullibility as widespread: people would be very often fooled into accepting empirically unfounded messages. Second, it views gullibility as often applying to costly beliefs, beliefs that lead to painful rituals, expensive purchases, risky rebellions, or harmful complacence. Third, it views gullibility as being mostly source-based: stemming from the undue influence of focal sources, often authority figures, be they religious leaders, demagogues, TV anchors, celebrities, and so forth. Most accusations of gullibility reviewed above share these traits. I will refer to this view of gullibility as strong gullibility."*

Mercier 2017

[110] *"Inconsistencies between background beliefs and novel information easily lead to belief updating. If John sees a green elephant in the yard, he updates his beliefs accordingly. John can afford to do this because his perceptual and inferential mechanisms do not attempt to mislead him. By contrast, in the case of communicated information, the honesty of the sender is open to question. This means that communicated information that is inconsistent with a receiver's background beliefs should be, on average, less likely to lead to belief revision than similar information obtained through perception (in the absence of contrary evidence provided by trust or arguments, see below, and Sperber et al., 2010). We would thus rely on plausibility checking, a mechanism that detects inconsistencies between background beliefs and communicated information, and that tends to reject communicated information when such inconsistencies emerge. ...*

There is substantial evidence that people detect inconsistencies between their background beliefs and communicated information, that such inconsistencies tend to lead to reject communicated information, and that information that is more inconsistent with one's prior beliefs is more likely to be rejected. ...

For instance, in a typical advice taking experiment, participants have to form an opinion about a given question—typically a numerical opinion, such as 'how much does the individual in this picture weighs?' They are then confronted with the opinion of another participant, and they can revise their opinion on this basis. When no relevant factor, such as expertise, differentiates the participant receiving the advice from the participant giving the advice, the advice tends to be heavily discounted in favor of the participant's own opinion (e.g., Yaniv & Kleinberger, 2000). Moreover, the advice is more discounted when it is further away from the participant's own opinion (Yaniv, 2004)."

Mercier 2017

[111] *"A major motivation for seeking advice is the need to improve judgment accuracy and the expectation that advice will help. An abundance of studies have shown that combining multiple sources of information improves estimation in the long run, in a variety of domains ranging from perceptual judgment to business forecasting (e.g., Armstrong, 2001; Sorkin, Hayes, & West, 2001; Yaniv, 1997). Aside from accuracy, there are also social reasons for seeking advice, which we consider only briefly here. Accountants performing complex audit tasks tend to solicit advice for self-presentational reasons and to increase the justification for their decisions (Kennedy, Kleinmuntz, & Peecher, 1997). Indeed, seeking advice implies sharing with others the responsibility for the outcome of a decision (Harvey & Fischer, 1997). One might argue, however, that even self-presentational reasons for seeking advice are rooted in the belief on the part of the individual or the organization that consulting someone else's opinion could improve one's final decision… A basic dilemma in using advice involves the amount of weight to place on others' opinions. Receiving advice often exposes decision makers to a potential conflict between their initial opinions and the advice. Consider a manager who believes that a certain new product is likely to gain success and is thus worthy of further development. The manager then receives a lukewarm expert opinion of her idea. How might she revise her opinion? The key question in many practical situations is to decide just how much weight ought to be assigned to a particular piece of advice. In particular, a decision maker's weighting policy might entail completely ignoring the other opinion, some adjustment of one's own opinion towards the other, or complete adoption of the other opinion... Previous work on the use of advice in decision making suggests a self/other effect whereby individuals tend to discount advice and favor their own opinion. In a judgmental estimation task (Yaniv & Kleinberger, 2000) respondents formed a final opinion on the basis of their initial opinion and a piece of advice. Rather than using equal weighting, respondents tended to place a higher weight on their own opinion than on the advisor's opinion. Even though the decision makers were sensitive to the quality of the advice (good vs poor), they tended to discount both good and poor advice. In a cue-learning study by Harvey and Fischer (1997), respondents shifted their estimates about 20–30% towards the advisor's estimates. Lim and O'Connor (1995) found that, in combining their prior personal forecasts and advisory (statistical) forecasts, judges weighted their own forecasts more heavily than the statistical forecasts."*

Yaniv 2004

[112] *"Receivers use a wide variety of cues to infer senders' trustworthiness. Some cues relate to the competence of the sender. A competent sender is a sender who is likely to have formed reliable beliefs. Cues to competence can be traits of senders such as dispositions (intelligence, diligence), or acquired expertise (being skilled in mechanics, say). Cues to competence can also be local, such as differences of perceptual access (who witnessed the crime).*

Other cues relate to the sender's benevolence. A benevolent sender is a sender who is likely to send messages that positively take into account the receivers' interests (Barber, 1983). Thus, benevolence entails more than the absence of lying. If a sender sends a message that only benefits the sender, and not the receiver, she would not be deemed benevolent, even if the message is not an outright lie. For instance, a friend who recommends a restaurant on the basis of preferences she knows not to be shared by her audience would not be very benevolent. Like cues to competence, cues to benevolence can be traits, stable features of senders that make them more likely to be benevolent toward the relevant receiver (relatedness to the receiver, say). Cues to benevolence can also be local. In particular, attention should be paid to how the interests of the sender are served by the acceptation of the message being evaluated. Self-interested messages should arouse suspicion.

...receivers take cues to trustworthiness into account in an overall sensible way, and that they lower their trust in senders who were committed to messages that proved unreliable… In line with this result, studies that have focused on informational conformity have found that it is taken into account, by adults and children, in a broadly rational manner : people tend to be more influenced by larger groups, by stronger majorities, and when they are less sure of their own opinions (Bernard, Harris, Terrier, & Clément, 2015; R. Bond, 2005; Campbell & Fairey,

1989; Gerard, Wilhelmy, & Conolley, 1968; McElreath et al., 2005; Morgan, Laland, & Harris, 2015; Morgan, Rendell, Ehn, Hoppitt, & Laland, 2012)."

Mercier 2017

[113] *"The studies mentioned above also show that the rejection of weak arguments does not stem from a blanket rejection of all arguments that challenge one's prior beliefs. Strong arguments are rated positively, and they influence participants, even when they run against prior preferences or beliefs (e.g., Petty & Cacioppo, 1979; Trouche et al., 2014; Trouche, Shao, & Mercier, 2017)...*

The experimental psychology results reviewed above demonstrate that people are endowed with mechanisms of epistemic vigilance that work, in the laboratory at least, reasonably well. They evaluate messages based on their content, on various attributes of their source, and on the arguments provided to support them: they are, broadly, ecologically rational."

Mercier 2017

[114] *"...most instances of gullibility are the outcome of content-based, rather than source-based, processes, and that they only indirectly bear on the working of epistemic vigilance."*

Mercier 2017

[115] *"Bloodletting is a salient example. One of the most common therapies for significant portions of Western history, inefficient at best and lethal at worse, it seems to have owed its success to the authority granted the writings of Galen and other prestigious physicians (Arika, 2007; Wootton, 2006). This thus seems to be a blatant example of misplaced prestige bias (Henrich & Gil-White, 2001). However, ethnographic data reveal that bloodletting is a common practice worldwide, occurring in many unrelated cultures, on all continents, including in many cultures which had not been in contact with Westerners (Miton, Claidière, & Mercier, 2015). These ethnographic data, as well as some experimental evidence, suggests that bloodletting owes its cultural success to its intuitiveness: in the absence of relevant medical knowledge, people find bloodletting to be an intuitive cure (Miton et al., 2015). If this explanation is correct, trust would flow in the other direction: instead of bloodletting being practiced because it is defended by prestigious physicians, it is because some physicians practiced and defended bloodletting that they became prestigious. This explanation could easily be extended to the most common forms of therapies in premodern cultures, which all aim at removing some supposedly bad element from the body (laxatives, emetics, sudation, see, Coury, 1967). Similarly, people would not refuse to vaccinate their children because they follow Jenny McCarthy or other prominent antivaxxers. Instead these figures would become popular because they attack a very counterintuitive therapy (Miton & Mercier, 2015). This phenomenon would thus be similar to that of political or religious leaders who are mostly deemed charismatic and prestigious because they endorse popular positions. In neither case would people be gullibly following prestigious leaders, instead they would simply be heeding messages they find appealing, and then conferring some prestige on those who defend them. The spread of misguided beliefs would thus mostly rest on content- based rather than source-based processes."*

Mercier 2017

[116] *"Richard Dawkins (1976) has proposed a biological metaphor that also assumes that ideas compete but that does not assume they compete solely based on truth. Dawkins pictured culture as being composed of many individual units (the cultural analogue of genes) that undergo variation, selection, and retention. As a label for this cultural gene equivalent, he proposed the term meme. Dawkins's memes do not compete solely on truth—consider annoying commercial jingles or a chain letter that threatens doom if it is not reproduced and spread (Dawkins, 1976). In this article we follow Dawkins in explaining how ideas propagate using a variation-selection-retention approach, so to acknowledge our theoretical approach and unit of analysis, we often use the term meme for ideas that propagate in the social environment."*

Heath et al. 2012

[117] *"Cultural evolution is a vibrant, interdisciplinary, and increasingly productive scientific framework that aims to provide a naturalistic and quantitative explanation of culture, in both human and non-human species (Mesoudi 2011; Richerson and Christiansen 2013). 'Culture' is commonly defined as the body of information that is transmitted from individual to individual via social learning (rather than genetically), and colloquially includes such phenomena as attitudes, beliefs, knowledge, skills, customs and institutions. Inspired by pre-existing population genetics tools, the mathematical models of cultural dynamics developed by Cavalli-Sforza and Feldman (1981) and Boyd and Richerson (1985) first established that cultural change can be modelled as*

an evolutionary process yet one that is not slavishly identical in its details to genetic evolution. Today, while maintaining a solid modelling core (e.g. Kendal et al. 2009; Rendell et al. 2010; Aoki et al. 2011; Lewis and Laland 2012; Aoki et al. 2012; Kempe et al. 2014), a wide range of methodologies are used in the field of cultural evolution, including phylogenetic analysis (e.g. Gray and Jordan 2000; Tehrani and Collard 2002; Lycett 2009; Currie et al. 2010; Tehrani 2013; O'Brien et al. 2014), laboratory experiments (e.g. Mesoudi et al. 2006; Caldwell and Millen 2008; Mesoudi and O'Brien 2008; Kirby et al. 2008; Morgan et al. 2012; Derex et al. 2013; Muthukrishna et al. 2014; Tamariz et al. 2014), ethnographic field studies (e.g. Guglielmino et al. 1995; Henrich and Henrich 2010; Mathew and Boyd 2011; Hewlett et al. 2011; Demps et al. 2012; Kline et al. 2013), quantitative analysis of pre-historical, historical, and contemporary datasets (e.g. Shennan and Wilkinson 2001; Henrich 2001; Kline and Boyd 2010; Collard et al. 2011; Turchin et al. 2013; Acerbi and Bentley 2014; Beheim et al. 2014), and comparative studies of culture across species (Whiten et al. 1999; Laland et al. 2011; Dean et al. 2012). Although varied in methodology and topic, these studies are united by the notion that culture evolves according to broadly Darwinian principles. In parallel with this approach, a group of cognitive anthropologists have advanced a similar project aiming towards naturalistic explanations of culture, mainly focusing on the role that cognitive factors play in the transmission and transformation of cultural representations (Sperber 1996; Atran 1998; Boyer 2001; Sperber and Hirschfeld 2004). This approach has generated findings using laboratory experiments (e.g. Boyer and Ramble 2001; Barrett and Nyhof 2001; Norenzayan et al. 2006; Fessler et al. 2014) and analyses of historical (e.g. Nichols 2002; Norenzayan et al. 2006; Morin 2013) and cross-cultural (e.g. Atran 1998) datasets. The two approaches initially developed separately and, despite a series of attempts at seeking common ground (Henrich and Boyd 2002; Claidière and Sperber 2007; Sperber and Claidière 2008; Henrich et al. 2008), there is remaining disagreement (see e.g. Claidière et al. 2014). This disagreement rests, at a general level, in a different view of cultural transmission. For the standard cultural evolution approach, typified by Boyd, Richerson, Henrich and others, it is common to think of cultural evolution as a process of selection between different variants (e.g. beliefs, ideas or artefacts) or models (referring to people from whom one can copy). When deciding a name for a newborn, for example, one chooses from a pool of variants—the existing names in the population—and the individual-level processes of selection determine the success, at the population-level, of the variants. Cultural transmission has relatively high fidelity, and selection between faithfully transmitted variants plays an important role in determining cultural trajectories. Sperber, Claidière, Atran, Boyer and colleagues, instead, argue that in the vast majority of cases cultural traits are neither properly copied or selected, but reconstructed each time an instance of transmission happens. The permanence of some cultural traits occurs not due to high fidelity cultural transmission but instead due to the existence of stable "cultural attractors" (Sperber 1996). For example, in an oral transmission of a story, say Cinderella, it is highly unlikely the story will be repeated verbatim at each passage. Still, some defining features, say the pumpkin coach or the wicked stepmother, perhaps because they are particularly memorable, will act as attractors, and will be repeated ('reconstructed') each time by different narrators. Cultural transmission here has relatively low fidelity, and non-random distortions and reconstructions play an important role in maintaining cultural diversity and stability.

This general divergence has a series of consequences, ranging from what are considered the most important or interesting factors to take into account when explaining the permanence and diffusion of cultural traits (cognitive transformation of representations for Sperber and colleagues, interaction of simple decision-making biases with populational dynamics for the standard cultural evolution approach) to how far the analogy between cultural and biological evolution should be pushed (less for the former than for the latter approach)."

Acerbi & Mesoudi 2015

[118] *"Many rumors convey information about potential danger, even when these dangers are very unlikely. In four studies, we examine whether micro-processes of cultural transmission explain the spread of threat-related information. Three studies using transmission chain protocols suggest a) that there is indeed a preference for the deliberate transmission of threat-related information over other material, b) that it is not caused by a general negativity or emotionality bias, and c) that it is not eliminated when threats are presented as very unlikely. A forced-choice study on similar material shows the same preference when participants have to select information to acquire rather than transmit. So the cultural success of threat-related material may be explained by transmission biases, rooted in evolved threat-detection and error-management systems, that affect both supply and demand of information. ..*

Threat-related information is a central theme of many rumors, as well as urban legends and religious myths (Allport & Postman, 1947; Boyer & Parren, 2015; Difonzo & Bordia, 2007; Fessler, Pisor, & Navarrete, 2014; Heath, Bell, & Steinberg, 2001; Stubbersfield, Tehrani, & Flynn, 2014). Previous studies of transmission shed

light on very general factors that may affect the cultural spread of particular kinds of material. For instance, people may want to transmit information with salient emotional content. Newspaper stories that become 'viral' in terms of Internet transmission, contain more emotional material than other stories (Berger & Milkman, 2012). Urban legends that elicit disgust are better transmitted than control items (Heath et al., 2001). Another related factor is a form of 'negativity bias' (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001) that has been observed in many domains of attention and memory (Ito, Larsen, Smith, & Cacioppo, 1998; Rozin & Royzman, 2001), as well as in the transmission of information (Heath, 1996). Negativity affects not just transmission but also belief, as people judge as more plausible the same information when it is framed in negative terms (e.g., "10% of heart transplants fail within a year") rather than with positive framing (e.g., "90% of heart transplants are fine after a year") (Fessler et al., 2014; Hilbig, 2009, 2012). However, negativity may be too broad a category. In functional terms, it is difficult to see why minds should be biased towards negative material in general. The bias reported may be more specific, as a focus on information related to potential danger, which makes more functional sense (Pratto & John, 1991). In four studies, we observed a privilege of threat-related information over other positive or neutral items, but also over negative material. In transmission chains, people preferred to transmit threat-related information (Study 1). Adding explicit information to the effect that the dangers described were unlikely did not modify this preference for threat- related items over negative items (Study 2). Even adding potentially more relevant positive information about the products left the threat- related information items as more likely to survive cultural transmission than any other material (Study 3). Finally, a forced-choice paradigm on these same materials showed that participants chose to seek more information about threat-items than about other aspects of the products described (Study 4).Taken together, these results support the hypothesis, that threat-related information is selected for social transmission, in preference to other kinds of information. As the less transmitted items include negative information, we should not interpret the advantage of threat-information as a form of negativity bias.

Blaine & Boyer 2017

[119] *"Overall, the analysis presented here suggests that one of the factors that could explain the success of online misinformation is that it appeals to general cognitive preferences. Consistent with previous research, 'suspect' articles were found heavily leaning towards negative content. The various cognitive factors coded were present to a different degree. Descriptions of threats were prominent, with almost 30% of the articles containing them. Elements eliciting disgust and sexual details were also present, but they were generally co-occurring with threat-related information (the single most successful 'fake news' in Facebook in 2017 is a good example of this combination: Babysitter transported to hospital after inserting a baby in her vagina, BuzzFeed, 2017). … Articles with minimally counterintuitive elements were less common than articles with threat-, sex-, and disgust-related information. In addition, violations of intuitions that could be considered 'supernatural' in the common sense of the term were even less, making for around 5% of the articles (the other articles consisted in violations of essentialist intuitions, see below). This is partly surprising, giving the importance given in the cultural evolution literature to MCI elements. … Social information and presence of celebrities were the elements quantitatively most important. …*

This analysis suggests a few general considerations on the spread of online misinformation. First, articles concerning political misinformation, while abundant, were still technically a minority in the sample considered (40% of the articles). Different sampling methods could, of course, give different results, but this figure is consistent with the idea that online misinformation is not necessarily political misinformation. While there may be good reasons to focus the attention to the possible risks that the spread of political 'fake news' online entails, it may also be conceivable that the danger of misinformation online has been overstated by previous research, by artificially limiting the breadth of the phenomenon on explicitly malicious political articles (similar conclusions on the overestimation of the effect of political misinformation are reached, for example, in Allcott and Gentzkow, 2017 and Guess et al. 2018)…

Second, is there any specificity of the spreading of online misinformation? Various reasons have been proposed to explain why misinformation should thrive online (as opposed to offline), including the fact that everybody can quickly and cheaply spread information, that digital media make easier to find other individuals confirming incorrect information, that online interactions can preserve anonymity (Allcott and Gentzkow, 2017) and that search engines, and especially social media algorithms, are optimised for (shallow) engagement, giving disproportionate weight to 'like' and previous traffic (Chakraborty et al. 2016). This analysis points to the fact that, however, the same features that make urban legends, fiction, and in fact any narrative, culturally attractive also operate for online misinformation. While this does not exclude that specific mechanisms favour online

spreading of misinformation, it suggests that to better understand them, some knowledge of why some narratives are attractive and others are not can be useful."

Acerbi 2019a

[120] *"All this show the paradox of junk culture in even starker relief. Psychologists have gathered large amounts of evidence for a series of cognitive systems to acquiring useful, that is fitness-relevant information about the world, especially from conspecifics, and ensuring that information is of sufficient quality. This seems to be a straightforward consequence of cognitive evolution. In the same way as our visual system is designed to use available information from light reflectance, our inference systems should be designed to acquire reliable information, as every increment in that capacity does translate as a survival advantage. So, again, why would humans blithely fill their minds with poor-quality information, which in most cases is of no clear advantage?"*

Boyer 2018

[121] *"This does not mean that plausibility checking or reasoning cannot fail in their own right. People could be tricked by sophistry and various other types of manipulative messages to accept messages that are in fact largely inconsistent with their beliefs (see, e.g., Maillat & Oswald, 2009; Stanley, 2015)."*

Mercier 2017

[122] *"Rumors are about mostly negative events and their sinister explanation. They describe people intent on harming us or who have already done so. They describe situations that will lead to disaster if no action is taken. The government is involved in terrorist attacks against the population, medical authorities conspire to spread mental illness in children, ethnic others are trying to invade us, and so forth. in other words, rumors describe potential danger and the many ways in which we could all be threatened. "… Human minds comprize specialize systems for threat detection. it is an evolutionary imperative for all complex organisms to detect potential dangers in their environment and engage in adequate precautionary behaviors… Threat-response systems, in humans as in other animals, face the problem that there is an important asymmetry between danger cues and safety cues. The former are actual properties of the environment… There is however no clear signal of non-danger… In humans, whose behavior is strongly affected by information from conspecifics this asymmetry of threat and safety has one important consequence, that precautionary advice is rarely put to the test. Indeed, it is one of the great advantages of cultural transmission that it spares individuals from systematically testing their environments to identify sources of danger… This would suggest that threat-related information is often considered credible, at least provisionally, as a precautionary measure. The psychologist Dan Fessler tested this directly by measuring people's acceptance of statements phrased in either negative, threat-related terms (such as "10% of heart attack patients die within 10 years") or positive terms ("90% of heart-attack patients survive for more than 10 years"). Ever though the statements are strictly equivalent, people place more confidence in the negatively framed ones. Similarly, people find the authors of descriptive texts, for example, about a computer program or a hiking trip; more competent and knowledgeable if the texts include threat-related information… So we should expect that people are particularly eager to acquire threat-related information. Naturally not all such information could give rise to rumors that people take more seriously than mere urban legends, otherwise cultural information would consist in nothing but precautionary advice. But several factors limit the spread of rumors about potential threats. First, all else being equal, plausible warnings have an advantage over descriptions of highly unlikely situations. … It is generally easier to convince our neighbors that the grocer sells rotten meat than that he occasionally turns into a reptile. Note that, as a matter of course, what is or is not plausible depends on the listener's own metrics… Second, the niche for non-tested (and generally invalid) precautionary information requires that the cost of precautions be relatively moderate. To take an extreme case, it is relatively easy to convince people not to walk around a cow seven times at dawn… Third, the potential cost of noncompliance, what would happen if we failed to take precautions, should be described as serious enough that the listener's threat-detection systems are activated… So it would seem that threat detection is one of the domains in which we may have to turn down our epistemic vigilance mechanisms and take as a guide to behavior precautionary information, especially of it is not too costly to follow, and if the averted danger is both serious and uncertain."*

Boyer 2018

[123] *"Deception may be adaptive, if you can exploit others, but then it becomes adaptive for others to develop the symmetrical weapon, the ability to see through deception. There is an equilibrium when capacities for deception and detection are roughly equivalent. But that equilibrium is unstable. Any organism that is slightly better than others at deception will gain an advantage, so that it will transmit its deceptive skills to its offsprings, until these*

skills become the population average. But then an increase in detection skills become adaptive, and in a similar way will gradually become the average. This kind of arms race between deception and detection is common in nature. In the case of human communication, the arms race consists in competition between the capacity to make one's utterances persuasive, on the one hand, and the ability to protect one's own beliefs from deception, on the other."

Boyer 2018

[124] *"There is also a motivation to transmit, without which many people would cultivate their own poor-value information but there would be no rumors, no junk culture.*

In many situations the transmission of low-value information is associated with strong emotions. People consider information about viruses and vaccinations and government conspiracies as terribly important. When they transmit information about such topics, people are not just eager to convey but also eager to convince. They do pay attention to their audiences' reactions, and they consider skepticism highly offensive. Doubt is attributed to all sorts of wicked motives… Why are the beliefs so intensely moralized? One obvious answer is that the moral value of broadcasting the information, and of accepting it, is a straightforward consequence of the information conveyed… A crucial part of our evolved psychology consists in capacities and motivations for efficient coalition management. So, when humans convey information that may persuade others to engage in specific actions, we should try to understand this in terms of coalition recruitment. That is to say, we should expect that an important part of the motivation here is indeed to persuade others to join in some collective action. … Roughly speaking, stating that someone's behavior is morally repugnant creates consensus more easily than claiming that the behavior results from incompetence. The latter could invite discussions of evidence and performance, more likely to dilute consensus than to strengthen it."

Boyer 2018

[125] *"To sum up, even from the earliest stages of cognitive development human minds seem to be designed to acquire useful knowledge about their environment. I must insist on the word 'useful'. We should not assume that human minds are designed to acquire true information about their natural and social environments. That is an important difference. Just because something is a fact does not mean that humans are equipped to find out about it. Conversely many of our intuitive expectations lead us to false beliefs. … So it is important to remember that the human mind is not always philosophically correct or scientifically accurate. The assumptions it contains may not be true, but they are useful. Usefulness, then, refers to selective pressure."*

Boyer 2018

[126] *"Human cognition is distinct in the degree to which it is shaped by cultural learning (i.e., information learned from others), in addition to individual experience and genetics. The cognitive mechanisms that shape and bias the acquisition of cultural information (which includes emotional responses, food preferences, cognitive and behavioral heuristics, etc.) have broad relevance to understanding human cognition and behavior… Several important lines of theoretical work have sketched cognitive foundations for our species' capacity for cumulative cultural learning. These accounts variously emphasize, for instance, intention reading and attention sharing (Tomasello, Carpenter & Behne, 2005), ostensive pedagogy (Csibra & Gergely, 2009; Gergely & Csibra, 2005), cognitive fluidity (Mithen, 1996) and mental time travel (Boyer, 2008). Among these, culture–gene coevolutionary (CGC) theories focus on and model the evolutionary dynamics facing an emerging cultural species, in particular the interaction of genetic and cultural inheritance systems (Boyd & Richerson, 1985) and the learning biases these interactions select for (Cavalli-Sforza & Feldman, 1981; Eriksson, Enquist & Ghirlanda, 2007; Kendal, Giraldeau & Laland, 2009; Mesoudi, 2009; Richerson & Boyd, 2005). By bringing together empirical evidence of human ancestral history and evolutionary models focused on understanding our capacities for cultural learning, CGC theories have derived predictions supported by a wide range of evidence from social psychology, economics, field studies and paleoarcheology (Galef & Whiskin, 2008a; Laland, 2004; Mesoudi, 2009; Powell, Shennan & Thomas, 2009; Richerson & Boyd, 2005; for a review, see Henrich & Henrich, 2007; Richerson & Boyd, 2005: chapter 2). CGC approaches suggest, among others things, that learners should be selective about who they attend to for the purpose of cultural learning. They specify a suite of hypotheses about which cultural learning strategies most effectively extract useful, adaptive information."*

Chudek & Heirich 2011

[127] *"By bringing together empirical evidence of human ancestral history and evolutionary models focused on understanding our capacities for cultural learning, CGC theories have derived predictions supported by a wide range of evidence from social psychology, economics, field studies and paleoarcheology (Galef & Whiskin,*

2008a; Laland, 2004; Mesoudi, 2009; Powell, Shennan & Thomas, 2009; Richerson & Boyd, 2005; for a review, see Henrich & Henrich, 2007; Richerson & Boyd, 2005: chapter 2). CGC approaches suggest, among others things, that learners should be selective about who they attend to for the purpose of cultural learning. They specify a suite of hypotheses about which cultural learning strategies most effectively extract useful, adaptive information… Strategies concerning from whom to learn are termed 'model biases.' Some individuals are just better in certain domains, or possess more relevant information, and it pays to learn from them. Alongside cues based on age, sex, health and dialect (cuing ethnicity), CGC specifies three candidates for evolved model biases: skill bias, success bias and prestige bias. 'Skill bias' entails learners selecting models by direct perception of their competence, which can be inaccurate (i.e., when it is not obvious to naive learners how to judge competence) and costly (i.e., when careful observation of many individuals is needed to gauge skill differentials). ' Success bias' entails learners selecting models by the accumulated symbols of their success, which can vary between societies; for instance, greater wealth, fancier ornamentation (Malinowski, 1922) or bigger yams (Kaberry, 1941). 'Prestige bias' entails learners preferring information from models to whom other learners have preferentially attended or deferred (Henrich & Gil-White, 2001). Prestige bias facilitates more accurate and rapid learning by capitalizing on others' knowledge about who is worthy of attention. Because others' preference for better-quality models is, for a cultural species, fairly reliable across generations and cultures, prestige bias, as an adaptation for exploiting this regularity, is a good candidate for a genetically evolved, cultural learning bias. Prestige bias, a technical coinage, does not denote prestige's usual English meaning (an acknowledged status difference); rather, it refers to learners' preference for inferring cultural information from whomever receives more attention and/or freely conferred deference from other learners. This difference results when learners prefer to attend to and hang around their more skillful peers, avoid attending to less skillful ones or, more likely, both. CGC's key prediction is that cues of differential attention and/or deference, alongside other sources of information about model quality, will be exploited by members of a cultural species because they reliably discriminate better from worse models across societies and epochs. Among adults, evidence for CGC's predicted biases (see Henrich & Henrich, 2007, for a review) has emerged from social psychology (Henrich & Gil-White, 2001), behavioral economics (Pingle & Day, 1996), experimental anthropology (Efferson, Lalive, Richerson, McElreath & Lubell, 2008), field studies (Rogers, 1995) and even corollaries in non-human animals (Galef & Whiskin, 2008a; Galef & Whiskin, 2008b; Horner, Proctor, Bonnie, Whiten & de Waal, 2010)."

Chudek & Heirich 2011

[128] *"The evolution of the human cultural capacity – that is, for intergenerationally stable, high fidelity, social transmission – created a new selective environment in which mutations improving the reproductive benefits of such transmission were favored. Our ancestral psychology evolved within physical and phylogenetic constraints) into an increasingly well-organized and specialized battery of biases jointly designed to extract reproductive benefit from the flow of socially transmitted information. Prestige processes emerge from this evolved social learning psychology. Cultural transmission is adaptive because it saves learners the costs of individual learning. Once some cultural transmission capacities exist, natural selection favors improved learning efficiencies, such as abilities to identify and preferentially copy models who are likely to possess better-than-average information. Moreover, selection will favor behaviors in the learner that lead to better learning environments, e.g., gaining greater frequency and intimacy of interaction with the model, plus his/her cooperation. Copiers thus evolve to provide all sorts of benefits i.e., 'deference') to targeted models in order to induce preferred models to grant greater access and cooperation. Such preferred models may be said to have prestige with respect to their 'clients' the copiers). The above implies that the most skilled/knowledgeable models will, on-average, end up with the biggest and most lavish clienteles, so the size and lavishness of a given model's clientele the prestige) provides a convenient and reliable proxy for that person's information quality. Thus, selection favors clients who initially pick their models on the basis of the current deference distribution, refining their assessments of relative model worth as information becomes available through both social and individual learning. This strategy confers a potentially dramatic adaptive savings in the start-up costs of rank-biased social learning. Finally, because high-quality information 'expertise,' 'performative skills,' 'wisdom,' 'knowledge') brings fitness-enhancing deferential clients, models have an extra incentive to outexcel each other."*

Heinrich & Gil-White 2001

[129] *"Participants watched an 'attentional cuing' clip, where two models received unequal bystander attention. In this cuing scene two bystanders stood between the models, attending to only one of them — the 'prestigious model.' This prestige cuing was followed by four 10-seconds (s) 'test' clips, where those two models*

demonstrated different behaviors, preferences and labels. In all test clips solitary models demonstrated their preference towards an object; then participants' own preferences toward those same stimuli were recorded. The order in which models appeared and the identity of the prestigious model were counter-balanced across participants...

Our findings provide support for the existence of a domain-sensitive prestige bias in children's learning: children's learning from cultural models was biased by the mere preferential attention of bystanders, particularly on activities similar to those the model had been engaging in when she received bystander attention. These strong effects from a minimal manipulation suggest that prestige bias may be a potent pressure on cultural evolution. As predicted (Henrich & Gil-White, 2001), we witnessed biased learning in different domains, including potentially costly dietary preferences…

With regard to cross-domain effects, Henrich and Gil-White (2001) predicted that prestigious individuals are 'influential, even beyond their domain of expertise.' Recent developmental research (Fawcett & Markson 2010) has indicated that 2-year-olds who know a model shares their preferences in one domain (food or television shows) will only imitate that model's preferences in that same domain, not the other. This suggests that mere 'similarity cues' may not be strongly influential beyond their domain of expertise. In our work with 'prestige cues', we witnessed an interesting domain-based asymmetry. Our subjects' food and drink preferences trended toward prestige-bias when they saw artifact-use cues (in study 1, the combined food and drink measures registered a significant effect after a prestige cue, but each measure independently did not); however, their artifact-use preferences trended away from prestige-bias after seeing food cues. This raises the interesting possibility that children's inferences about model quality exploit an asymmetric map of the relationships between learning domains."

Chudek et al. 2011

[130] *"The evidence reviewed in this article provides mixed support for the use of prestige-biased social learning in both human adults and children. However, few studies have examined this and further research is needed to clarify which factors lead to variation in the use of prestige-biased social learning. The difficulty of the task, the relevance of the domain for the individuals and the benefits and costs associated with the task seem to be important factors influencing the use of prestige-biased social learning (see prediction (viii) in Table 2). In general, easy tasks, tasks that are not relevant for participants and tasks that do not provide incentives to perform well or avoid costs (e.g., monetary rewards or costs) seem not to stimulate the use of prestige-biased social learning (Acerbi and Tehrani, 2018). Other factors taken into account in the literature, such as experience and age (Little et al., 2015) seem to be important when they affect task difficulty, the relevance of the domain for the participants and potential gains or costs of the task for the participants. For instance, expertise leads to a greater use of prestige-biased social learning when the task is more relevant for the experts but the task is still difficult for them (Verpooten and Dewitte, 2017). Similarly, younger individuals use more prestige-biased social learning than older individuals when the task is more relevant for them (Little et al., 2015). Moreover, when there is little variation in knowledge/skill in a group, it is more adaptive to learn from low access cost models than from costly prestigious models (Henrich and Henrich, 2010; see prediction (ix) in Table 2). Another factor that influences the use of prestige-biased social learning is the availability of alternative social learning biases, e.g., success or content biases. When success information is provided, this information should be preferentially used over prestige information (prediction (ix) in Table 1). However, this was not found in the sole experiment comparing prestige with success bias (Atkisson et al., 2012), although this is a single study. Both direct and conceptual replications are needed to gain confidence in this result. Content bias was stronger than prestige bias in another study (Acerbi and Tehrani, 2018), but this might depend on the domain and the factors mentioned above (i.e., task difficulty, relevance for the individual, and benefits and costs associated with the task). Variation in some of these factors (e.g., the relevance for the participants) might lead some participants (e.g., non-experts) to make use of content biases, while other participants (e.g., experts) to employ prestige-biased social learning (Verpooten and Dewitte, 2017). It is also possible that prestige biased social learning has different effects on different measures of influence, e.g., recall, likability, behavioural influence, willingness to transmit and receive information. To the extent these measures of influence affect task difficulty, relevance for the participants or the benefits/costs associated with tasks, it seems plausible that the different measures would be a source of variation in the use of prestige and other social learning biases. For instance, although one recent study found that anti-vaccination messages are not better transmitted per se, exploratory analyses showed that when anti-vaccination messages are provided by doctors (i.e., a prestigious source within a relevant domain) these types of messages are especially powerful in influencing people's vaccination-related decisions (Jiménez et al., 2018). Similarly, although people might be able to appreciate the*

content of certain pieces of information (e.g., quotes, news, artworks, etc.) independent of the prestige of the source of the information, they might be more influenced by prestige cues when they want to achieve influence over other people's behaviour (e.g., by quoting a prestigious source of information), get personal or social benefits (e.g., choosing artworks to be displayed in their own town) or they have to decide whether to learn more about a topic or transmit the information about the topic to other people. Therefore, research on prestige-biased social learning might benefit from comparing the influence of prestige cues on different types of outcomes…

In conclusion, H&GW's theory of the evolution of prestige has generated a great deal of research and this research has stimulated new research questions and predictions. Although the evidence reviewed here suggest that prestige-bias social learning is employed in at least some contexts, further research will need to determine the precise circumstances in which people use prestige cues to learn socially, and when the use of these cues is adaptive."

Jimenez & Messoudi 2019

[131] *"In our studies we presented respondents with questions that had real consequences for them as decision makers, since they received a bonus for making accurate judgments. The respondents were given advice and the principal measure was the weight placed on that advice in their final decisions. The studies, which were conducted on a computer due to their interactive nature, shared the following general procedure. In the first phase, respondents were presented with questions and asked to state their estimates. In the second phase, they were presented with the same questions along with estimates made by various advisors (other students). The respondents were then asked to provide their estimates once again. They were free to use the advice as they wished… A coherent picture emerges from the advice weighting policies observed across the studies. First, the results of Study 1 show egocentric discounting of advice. Second, advice discounting was not indiscriminate; individuals had a veridical view of their knowledge, so that the less knowledgeable ones placed greater weight on the advice (Studies 1–3). Third, the weight of advice declined with the distance between the advice and their initial opinions (Studies 2–3); this distance effect was exhibited in the high-knowledge condition and to a lesser extent in the low-knowledge condition as well. … The asymmetric weighting of one's own and other opinions is attributed to the fundamental asymmetry in access to the underlying justifications for each opinion. Decision makers can assess what they know and the strength of their own opinions, but are far less able to assess what an advisor knows and the reasons underlying her/his opinions. Naturally, one's confidence about a given opinion (or hypothesis) is related to the amount of evidence that one could readily recruit to support it. Other things being equal, decision makers are likely to feel more confident about their own opinion than about the other opinion, hence their own estimate would receive greater weight than the advice. Earlier findings suggest that respondents weight each opinion according to the expertise ascribed to its source (Birnbaum & Stegner, 1979; Birnbaum & Mellers, 1983). The self/ other asymmetry presumably enhances the expertise ascribed to the self. This line of reasoning about information asymmetry is also reminiscent of the principal- agent problem in organizations (Eisenhardt, 1989).*

The explanation of the self/other effect in terms of differential information access seems preferable to alternative explanations that posit either a self-serving bias (e.g., an optimistic bias) or commitment to one's past decisions as the root of discounting others' views. To be sure, self-serving biases pervade interpersonal comparisons, in that, for example, people believe that they have lower chances of experiencing negative life events, such as car accidents and strokes, than others do or that they rank higher than others on various abilities and attributes, such as driving ability and social skills (e.g., Brown, 1986). But a bias of this sort does not readily explain respondents' weighting policies for advice, especially the sensitivity of those policies to the respondents' own knowledge (Studies 1–3) and their sensitivity to the quality of the advice (Yaniv & Klein- berger, 2000).

Commitment to one's past decisions is a powerful motive in decision making, yet it cannot readily explain the findings either. The antecedents of commitment—high costs for being inconsistent, the need to justify decisions to others, having to admit past mistakes, and having to save face with respect to ego-involving issues—were largely absent in the present studies. Our respondents made their judgments in a private setting (by entering responses into a computer file), received incentives for accuracy, and were not asked to justify their estimates.

A cognitive explanation based on informational asymmetry and the assessment of available evidence is more parsimonious and hence superior to those based on a self-serving bias or commitment because it can readily account for the finding that respondents' weights on advice are sensitive to the quality of the advice (Yaniv & Kleinberger, 2000) as well as their own knowledge (e.g., Study 1), without making unnecessary assumptions."

Yaniv 2004

[132] *"The literature on social cognition emphasizes infants' emerging understanding of other people as repositories of knowledge who can adopt varying cognitive and emotional perspectives toward objects and events (e.g., Baldwin & Moses, 1996; Moses, Baldwin, Rosicky, & Tidball, 2001; Tamis-LeMonda & Adolph, 2005). With the realization that caregivers can provide useful social information, infants increasingly turn to them for guidance when they are unsure about how to respond. Infants' capacity to benefit from others' social advice enormously expands opportunities for learning because infants need not rely solely on learning through self-discovery; now infants can seek others' advice in ambiguous situations and respond to unsolicited advice about how to act (Moses et al., 2001)."*

Tamis-LaMonda et al. 2008

[133] *"Cultural evolution theory posits that a major factor in human ecological success is our high-fidelity and selective social learning, which permits the accumulation of adaptive knowledge and skills over successive generations. One way to acquire adaptive social information is by preferentially copying competent individuals within a valuable domain (success bias)."*

Jimenez & Messoudi 2019

[134] *"Cultural evolutionary theory suggests that individuals should be selective with respect to when they adopt the decisions of others (Boyd & Richerson, 1985; Rogers, 1988), and that natural selection will lead to the use of adaptive learning strategies that guide the use of social information (Boyd & Richerson, 1985; Henrich & McElreath, 2003; Laland, 2004). Such 'social learning strategies' (also known as 'transmission biases'; Boyd & Richerson, 1985) have been primarily examined through population genetic and game theory modeling (Cavalli-Sforza & Feldman, 1981; Boyd & Richerson, 1985; Rogers, 1988; Feldman, Aoki & Kumm, 1996; Schlag, 1998, 1999; Wakano & Aoki, 2007; Enquist, Eriksson & Ghirlanda, 2007; Kendal, Giraldeau & Laland, 2009; Nakahashi, Wakano & Henrich, 2012; Kandler & Laland, 2013), and through experiments with human adults (McElreath, Lubell, Richerson, Waring, Baum, Edsten, Efferson & Paciotti, 2005; Efferson, Lalive, Richerson, McElreath & Lubell, 2008; Mesoudi, 2008, 2011; Toelch, Van Delft, Bruce, Donders, Meeus & Reader, 2009; Toelch, Bruce, Meeus & Reader, 2010; Morgan, Rendell, Ehn, Hoppitt & Laland, 2011). One such bias – a tendency to copy others when uncertain as to how to solve the task at hand – has been a central assumption of theoretical models of cultural evolution.*

Boyd and Richerson (1988) modeled individuals in a spatially and temporally variable environment. They postulated than when asocially acquired information left individuals uncertain, they should adopt the decisions of others. Enquist et al. (2007) considered a related rule called 'conditional social learning', by which individuals first learn asocially, but go on to learn socially if the result of their asocial learning is unsatisfactory, an outcome that is likely on more difficult tasks. Their analysis found this rule to be a successful strategy across a range of conditions – particularly when asocial learning is relatively cheap (i.e. energetically undemanding and/or low risk) (Enquist et al., 2007).

Evidence for a bias to copy others when uncertain also comes from empirical studies with non-human animals (Van Bergen, Coolen & Laland, 2004; Galef & Whiskin, 2008). In adult humans, across multiple tasks, individuals' confidence ratings in their performance strongly predicted whether they would revise their decision when presented with conflicting social information (Morgan et al., 2011; See Morrison, Rothman & Soll, 2011; Soll & Mannes, 2011; Minson & Mueller, 2012). Furthermore, individual confidence ratings were shown to predict accuracy, supporting the notion that this strategy increases performance (Morgan et al., 2011)."

Morgan et al 2014

[135] *"Metacognition, the ability to think about our own thoughts, is a fundamental component of our mental life and is involved in memory, learning, planning and decision-making."*

Grimaldi, Lau & Basso. 2015

[136] *« La cognition désigne l'ensemble des traitements de l'information effectués par le cerveau, notamment ceux qui sont mis en jeu dans les apprentissages scolaires. La métacognition désigne les processus cognitifs qui contrôlent et évaluent la cognition elle-même. Autrement dit, le terme de "métacognition" désigne l'ensemble des processus par lesquels chacun d'entre nous régule son attention, choisit de s'informer, de planifier, de résoudre un problème, repère ses erreurs et les corrige. À l'école, cet ensemble de capacités joue un rôle central. Une "bonne régulation" conduit l'élève à s'engager dans l'apprentissage avec confiance et enthousiasme. La "mauvaise régulation" de la métacognition se solde par le dégoût d'apprendre, l'évitement de l'école, le décrochage, et par ce que l'on nomme "la spirale de l'échec".* »

Proust 2019

[137] "*Metacognition is a broad term, and often interpreted differently by different researchers. As a first step, it is crucial to separate the empirical definition of metacognition from its epistemological status as a meta-level representation of an object-level cognition. Empirically, metacognition is often operationalized as 'behaviour about behaviour' rather than 'cognition about cognition' (see table 1 in Fleming & Dolan). Here, we define second-order behaviours as decisions contingent on other behavioural outputs (that either have occurred or will occur). Consider a visual detection task. Following a first-order response as to whether the stimulus is present or absent, a confidence judgement in one's response being correct is second-order with respect to the previous decision. This does not necessarily entail that the second-order judgement requires a meta-level representation of the object-level decision; it could instead be accomplished via object-level representations, for example, by basing confidence on information about the stimulus. Alternatively, the confidence judgement could be based on a meta-representation of the decision and subsequent response. This creates an initial division of the theoretical landscape, with two orthogonal dimensions—those of level of representation and order of behaviour.*"

Fleming Dolan & Frith 2012

[138] "*The function of metacognition has usually been construed in terms of control of behaviour and mental processes: the implication is that accurate control requires accurate monitoring. However, recent attempts to distil metacognition from other potentially confounding variables have led to a seeming paradox: if, to measure metacognition, one needs first to discount the influence of first-order behaviour, then the functional benefits of metacognitive capacity would appear moot. Perhaps, instead, metacognitive capacity is not necessary for controlling simple behaviours, but becomes relevant for complex abilities such as reasoning and planning. Fletcher & Carruthers critically appraise this hypothesis, presenting a deflationary account of a metacognitive system for human reasoning. They argue that reasoning can often be explained by an appeal to individual strategies acquired through individual and cultural learning. However, they also suggest that a dedicated 'mind-reading system' is recruited in a self-monitoring capacity to guide reasoning, an example of a system originally evolved to serve social functions playing a special role in self-directed cognition (see Carruthers). Frith has proposed that metacognitive states are useful precisely because they can be communicated to others, promoting social interaction. Bahramiet al. further suggest that communicating meta-cognitive confidence may act to replace explicit feedback about decision outcomes, and thus provide an ecologically relevant role for metacognition in social learning… Finally, an intriguing idea is that metacognition is functional precisely because it enables representation of the absence of knowledge. Object-level representations are often concerned with presence of stimuli in the world; they rarely deal in absence (consider the failure to represent the blindspot in vision). In contrast, 'knowing I do not know' is a meta-level representation of the absence of object-level memory. Investigating this putative function may benefit from greater integration with work quantifying epistemic behaviour—by sampling information over time, an agent can adaptively reduce its uncertainty, achieving a balance between the additional cost of exploration and the benefit of gaining further information. Interestingly, one key variable in driving further information-gathering is the current (im)precision of one's representation, providing a connection to computational models of confidence discussed in §4. These suggestions for functional roles are not mutually exclusive.*"

Fleming, Dolan & Frith 2012

[139] « *Qu'il s'agisse d'une action physique ou cognitive, le cerveau ne nous laisse pas nous engager dans une activité sans avoir anticipé nos chances de réussite. Les travaux expérimentaux démontrent que la première source d'information utilisée par l'auto-évaluation métacognitive consiste dans les prédictions de succès ou d'échec fournies par les sentiments qui sont produits au cours de l'activité. Ils sont très différents selon l'activité cognitive et le segment de l'activité qu'ils concernent (par exemple sa faisabilité à venir ou sa correction finale).* »

Proust 2019

[140] « *Comment le cerveau sélectionne-t-il ses indices prédictifs, c'est-à-dire le critère de réussite pour une activité donnée ? Il les choisit par renforcement. Grâce aux observations de l'activité neurale faites sur l'animal, on connaît maintenant quelques-uns des indices prédictifs utilisés par le cerveau. Par exemple, la rapidité avec laquelle le cerveau commence à activer le traitement de l'information liée à la tâche est un indicateur prédictif. D'autres indicateurs sont liés à la manière dont s'effectue la coordination entre les assemblées neuronales, en particulier leur rapidité de convergence vers une seule décision… Dans les deux cas, les indices prédisant le succès – par exemple : la rapidité de l'activation neuronale ou la rapidité de la convergence vers une seule décision – sont comparés aux indices observés dans le présent contexte. L'ensemble*

des indices forme ce qu'on appelle des "heuristiques prédictives" (voir figure 4). Nous ne savons pas quelles sont les heuristiques que notre cerveau a utilisées, mais nous éprouvons les sentiments qui en résultent. Ce sont ces sentiments qui nous permettent, à chaque instant, de savoir ce que nous savons, ce que nous comprenons, ce que nous avons envie d'apprendre, et ce que nous avons réussi ou échoué à faire. »

Proust 2019

[141] "*… Hampton (Hampton, 2001) devised a prospective memory confidence task. Trained monkeys performed a delayed-match-to-sample task. In this task, an image, referred to as the target, appeared at the beginning of a trial. At the end of the trial, after a delay, animals were required to select the target that reappeared with another series of images (distractors). To evaluate if the monkeys remembered the target, after two thirds of the delay, the monkeys received an option to accept the test or decline it. If they accepted it, and they made a correct match, they received a large reward. However, if they made a mistake by choosing a distractor as a match, then they received no reward. If the monkeys declined the test, they received a small reward, regardless of whether they chose the target or a distractor as the match. The investigators reasoned that if the monkeys believed they would perform well, they would accept the test, choose the correct target and receive a large reward. However, if they were uncertain, they would decline the test, and opt for the small but certain reward. Therefore, in this task, the monkeys made a prospective judgment about how they were likely to perform on the test. An additional strength of the task design was that four stimuli were used as possible targets and were selected as targets randomly each day. Because the stimulus set changed across sessions, the monkeys could not associate one particular stimulus with the likelihood of correct responding. Rather, they had to rely on their memory of the sample stimulus to decide whether to accept or decline the next step of the task. In addition, the monkeys could not use cues such as their own reaction time to estimate the likelihood of a correct response, because they had to decide to take the test or not before the match choice was required. As predicted, monkeys opted out when they did not remember, consistent with the hypothesis that they were less confident on those trials. Indeed, they performed better in this task then in a similar forced task, when they did not have the option to decline the test.*

In a similar spirit, Son and Kornell (2005) trained rhesus macaque monkeys to distinguish the length of two lines. After the monkeys made their decision, consisting of choosing the longest line, they were required to rate their confidence in their decision by making a bet, that is, a retrospective task. Two betting options were represented by two choice targets. If the monkeys chose the low bet target, they received a small reward, regardless of whether their previous response on the discrimination task was right or wrong. If they chose the high bet target, they received a large reward for correct responses and no reward for incorrect responses. Monkeys generally chose low rewards more frequently in difficult discrimination trials indicating that they knew when they did not know. The same monkeys engaged in the same betting strategy during a dot-density discrimination task, showing that they could generalize their reports of confidence to different tasks. Similar approaches have been used to study confidence in smaller mammals such as rodents. Foothe and Crystal (2007) trained rats to discriminate the duration of sounds. In each trial, the rats were able to choose if they wanted to take a test or not. Similar to the monkeys, rats chose to avoid the test when the stimulus was ambiguous."

Grimaldi, Lau & Basso 2015

[142] "*Often, researchers restrict the definition of metacognition to the kind of second-order behaviour available for subjective report [6]. In many of these cases, it seems intuitive that the report is capturing an aspect of cognition that is secondary to the cognitive process itself. Take the case of blindsight [9]: in some patients with lesions to primary visual cortex, visuomotor performance when responding to targets in the 'blind' field may be well above chance, yet the patient reports not seeing anything. This is a case where first-order (visuomotor) performance is high, but awareness is absent. Yet, the reliance on subjective reports to index metacognition precludes the ascription of metacognition to non-human animals and non-verbal infants, and may prematurely equate metacognition with consciousness (see §6). In contrast, non-verbal behavioural measures do not suffer from these drawbacks. Smith et al. review a large body of work in non-human animals using the 'uncertain-option' paradigm… whether the second-order behaviours often used to index metacognition can be explicable in non-metarepresentational terms remains an empirical question. We might find that an object- level account is sufficient to explain second-order behaviour in some circumstances, but not others. On the other hand, evidence from human neuropsychology that first- and second-order behavioural performances are dissociable suggests that at least some degree of separate representation will be required to account for second-order behaviours."*

Fleming, Dolan & Frith 2012

[143] "*Converging theoretical and empirical research suggests that most animals implicitly represent properties of their cognitive processes and use these for cognitive control … This challenges the view that only humans have*

metacognition .., which is broadly defined as 'cognition about cognition', and raises the question why humans, unlike other animals, not only implicitly but also explicitly represent properties of their cognitive processes…

"There is compelling evidence that non-human animals are more likely to seek additional information … to opt out of making … and to make lower post-decision wagers … under conditions in which a human observer would describe them as uncertain; for example, when the animal is required to make a difficult rather than an easy visual discrimination, or to remember an event over a long rather than a short interval. Some recent studies of monkeys … rats … and pigeons … have also indicated, using transfer tests and single neuron recording, that this type of metacognitive behaviour can be regulated by internal rather than external cues; for example, that it covaries more precisely with neural signals from the orbitofrontal cortex or the supplementary eye fields than with external stimulus values."

Shea et al. 2014

[144] *"Metacognitive confidence can be formalized as a probability judgment directed toward one's own actions—the probability of a previous judgment being correct. There is a rich literature on the correspondence between subjective judgments of probability and the reality to which those judgments correspond. For example, a weather forecaster may make several predictions of the chance of rain throughout the year; if the average prediction (e.g., 60%) ends up matching the frequency of rainy days in the long run we can say that the forecaster is well calibrated. In this framework metacognition has a normative interpretation as the accuracy of a probability judgment about one's own performance."*

Fleming & Lau 2014

[145] *"Prospective metacognition refers to making judgments or predictions about what information will be available in memory in the future. Retrospective metacognition, in contrast, involves making judgments about a past experience, specifically about whether a memory item has been successfully encoded."*

Grimaldi, Lau & Basso 2015

[146] *"In studies of human perceptual decision-making, confidence is often measured with retrospective judgment. Subjects give a confidence rating right after a report about a perceptual experience and therefore must base their confidence judgment on the memory of their initial response. For example, a subject might first perform some perceptual task such as reporting their perception of an ambiguous object (do you see a vase or a face?). Then the subject would immediately declare how confident s/he felt about that decision.*

Similar to measures of confidence using open-ended ratings, several scales have been developed to measure confidence more quantitatively. The most commonly used is confidence rating. In this scale, the subject is asked to report confidence on a continuous scale ranging from 0% or complete uncertainty to 100% or complete certainty. Alternatively, it can be assessed with discrete fixed levels, or a simple binary choice (confident/not confident, Cheesman and Merikle, 1986; Dienes and Perner, 1999). However the use of ratings has been criticized because some subjects may find it not intuitive or they may be poorly motivated to accurately report their confidence (Persaud et al., 2007). To overcome these limitations, post-decision wagering has been introduced, in which subjects bet money or tokens on their own decisions (Persaud et al., 2007; Ruffman et al., 2001). In this context, subjects should ideally bet low when they are not confident and bet high when they are confident, in order to maximize gain. This task is more engaging and more intuitive for most participants. However, it has been noted that wagering can be influenced by individual propensity to risk (Fleming and Dolan, 2010) and that subjects tend to use only the ends of the scale, probably in order to maximize gains (Sandberg et al., 2010), thus suffering from low sensitivity for intermediate ranges. In an attempt to develop a scale that has both the sensitivity of confidence ratings and the intuitiveness of post decision wagering, the feeling of warmth scale has been developed (Metcalfe, 1986; Wierzchon et al., 2012). In this scale subjects report their confidence as a temperature, ranging from cold (not confident) to hot (very confident), with intermediate options (e.g. chilly or warm). The perceptual awareness scale (Ramsøy and Overgaard, 2004) and the Sergent-Dehaene scale (Sergent and Dehaene, 2004) are also commonly used and were developed to judge the degree of visibility in visual tasks, ranging from no visibility at all to clear perception, with discrete intermediate levels (perceptual awareness scale), or a continuous spectrum (Sergent- Dehaene scale). When applied to confidence, however, these two scales end up being very similar to confidence rating. An extensive discussion of the properties and sensitivities of the different scales is beyond the scope of this review. For a rigorous comparison see, Sandberg et al., (2010) and Wierzchon et al., (2012)."

Grimaldi, Lau & Basso. 2015

[147] "*Hampton (Hampton, 2001) devised a prospective memory confidence task. Trained monkeys performed a delayed-match-to-sample task. In this task, an image, referred to as the target, appeared at the beginning of a trial. At the end of the trial, after a delay, animals were required to select the target that reappeared with another series of images (distractors). To evaluate if the monkeys remembered the target, after two thirds of the delay, the monkeys received an option to accept the test or decline it. If they accepted it, and they made a correct match, they received a large reward. However, if they made a mistake by choosing a distractor as a match, then they received no reward. If the monkeys declined the test, they received a small reward, regardless of whether they chose the target or a distractor as the match. The investigators reasoned that if the monkeys believed they would perform well, they would accept the test, choose the correct target and receive a large reward. However, if they were uncertain, they would decline the test, and opt for the small but certain reward. Therefore, in this task, the monkeys made a prospective judgment about how they were likely to perform on the test. An additional strength of the task design was that four stimuli were used as possible targets and were selected as targets randomly each day. Because the stimulus set changed across sessions, the monkeys could not associate one particular stimulus with the likelihood of correct responding. Rather, they had to rely on their memory of the sample stimulus to decide whether to accept or decline the next step of the task. In addition, the monkeys could not use cues such as their own reaction time to estimate the likelihood of a correct response, because they had to decide to take the test or not before the match choice was required. As predicted, monkeys opted out when they did not remember, consistent with the hypothesis that they were less confident on those trials. Indeed, they performed better in this task then in a similar forced task, when they did not have the option to decline the test.*

In a similar spirit, Son and Kornell (2005) trained rhesus macaque monkeys to distinguish the length of two lines. After the monkeys made their decision, consisting of choosing the longest line, they were required to rate their confidence in their decision by making a bet, that is, a retrospective task. Two betting options were represented by two choice targets. If the monkeys chose the low bet target, they received a small reward, regardless of whether their previous response on the discrimination task was right or wrong. If they chose the high bet target, they received a large reward for correct responses and no reward for incorrect responses. Monkeys generally chose low rewards more frequently in difficult discrimination trials indicating that they knew when they did not know. The same monkeys engaged in the same betting strategy during a dot-density discrimination task, showing that they could generalize their reports of confidence to different tasks. Similar approaches have been used to study confidence in smaller mammals such as rodents. Foothe and Crystal (2007) trained rats to discriminate the duration of sounds. In each trial, the rats were able to choose if they wanted to take a test or not. Similar to the monkeys, rats chose to avoid the test when the stimulus was ambiguous.

Grimaldi, Lau & Basso 2015

[148] "*To address this issue, we combined a nonverbal memory-monitoring paradigm developed for rhesus monkeys (22) with a pointing paradigm suitable for human infants. Twenty-month-old infants (n = 80) had to remember the location of a hidden toy for a variable delay before pointing to indicate where they wanted to recover it (Fig. 1A). Task difficulty was manipulated along two orthogonal dimensions: (i) Infants had to memorize the location of the toy for a variable delay (3, 6, 9, or 12 s), and (ii) they either saw the toy being hidden at a given location (possible trials) or could not see where the toy was being hidden (impossible trials). Crucially, half of the participants were given the possibility to avoid responding by asking their caregiver for help (AFH) instead of pointing (experimental group; n = 40), whereas the other half were not given this opportunity and could only choose a location by themselves (control group; n = 40). This manipulation enabled us to test whether infants can monitor and communicate their own uncertainty. Indeed, if infants can monitor their own knowledge state, they should use the AFH option (i.e., opt-out) when they have forgotten the toy location, thereby avoiding mistakes and improving their performance …Furthermore, if infants can monitor the strength of their memory trace, they should use the AFH option more often at higher levels of uncertainty (i.e., for longer delays and impossible trials)… We then tested whether task difficulty had an impact on the probability of asking for help. Indeed, if infants were monitoring their own uncertainty about the toy location, they should have asked for help more often as the memorization delay increased. This analysis was restricted to the participants in the experimental group, who asked for help in at least one trial per condition (n = 21). An ANOVA revealed that the probability of asking for help was higher for impossible than for possible trials [Fig. 2A; F(1,20) = 24.22; P < 0.001]. Furthermore, within possible trials, the probability of producing an AFH response increased with increasing delays [Fig. 2B; F(1,20) = 4.62; P < 0.05]. Thus, infants' tendency to ask for help varied with task difficulty, suggesting that infants used the AFH option strategically to avoid responding when they felt uncertain about the toy location…*

When given the opportunity to decide whether they should respond by themselves or avoid responding by asking for help, 20-mo-olds are able to strategically adapt their behavior. That is, they selectively seek help to avoid making errors and to avoid difficult choices. In the comparative literature, these adaptive "opt-out" behaviors have been taken as evidence for metacognitive uncertainty monitoring in several species … However, some authors have argued that such behavioral patterns could also be explained by associative or reinforcement learning mechanisms … For instance, they suggest that difficult trials are simply avoided because individuals learn that the probability of obtaining a reward is lower for those trials … Whether or not this associative interpretation can be ruled out in comparative research, in which animals are extensively trained, remains a controversial issue … However, in the present study, an associative account seems unwarranted because infants only received a few trials (i.e., a maximum of two trials for each level of task difficulty), leaving little room for associative learning. Moreover, the proportion of AFH responses did not increase across time [effect of trial rank on the proportion of AFH responses: $F(1,20) = 0.22$; $P > 0.6$], ruling out an associative interpretation in terms of reinforcement learning. Another issue raised in the comparative literature concerns the fact that when the opt-out alternative is available simultaneously with another choice, some competition might take place between these options .... This might eventually lead to the opt-out option being triggered by default whenever the participant is unable to accumulate enough evidence and commit to a decision before a deadline has been reached. Under this account, infants in our study would simply ask for help by default when no memory is available to trigger an appropriate motor plan. However, if infants simply turned to their parents automatically when no response came to their mind (e.g., to seek comfort), we should observe a similar tendency in the control group. In fact, although infants in the control group were not taught that they could ask for help, and even though their caregiver remained unresponsive, we did observe a few spontaneous 'AFH-like' responses in this group [mean number of AFH responses in the control group: 0.6; in the experimental group: 1.42; $t(39) = 3$; $P < 0.005$; Fig. S3]. However, when we analyzed the frequency at which infants looked toward the parent in the control group, we found absolutely no increase with task difficulty (Fig. S3A), and excluding those trials did not impact performance (Fig. S3B). Thus, infants in the control group did not orient selectively toward their parents when they were more likely to have forgotten the toy location. In turn, this finding confirms that infants in the experimental group did not automatically turn toward their parents when no response came to their mind. Rather, our results are consistent with the idea that infants in the experimental group learned that they could communicate with their caregiver to obtain some help whenever they felt that they were likely to make an error."

Goupil, Romand-Monnier & Kouider 2016

[149] *"In experiment 1, 18-month-old infants (n = 29) saw an object being hidden in one of two opaque boxes and, after a delay, were asked to point to indicate where the object was concealed (see Figure 1A and the Supplemental Experimental Procedures). First-order performance on this task was assessed along a parametric variation of difficulty (i.e., memorizing the location of the toy for a brief or longer delay). Immediately following this choice, infants were provided with the selected box. The amount of time they were willing to search within this box before giving up was used as a measure of post-decision persistence. Importantly, persistence times (PTs) were measured in the absence of any external feedback on performance, allowing us to use this measure as a proxy for confidence [7]... Consistent with our hypothesis, we observed that infants searched longer in the box following correct as compared to incorrect decisions ($t(28) = 2.1$; $p < 0.05$)."*

Goupil & Kouider 2016

[150] *"The two boxes contained a lid, which could only be opened by an adult. Although one of them had a slit such that infants were able to directly reach for its content (unsealed box), there was no slit in the other box, rendering its content unreachable without the help of an adult (sealed box). Importantly, both boxes looked identical from the infants' point of view, such that they could not know which box they were selecting (sealed versus unsealed box). During the familiarization phase (four trials), infants saw the experimenter hide a biscuit in one of two boxes. The experimenter then asked them to point to indicate where they remembered the biscuit to be. As soon as infants produced a pointing response, the selected box was pushed forward. During the first two trials, the biscuit was hidden in the unsealed box, so infants could directly recover it. During the last two trials, the biscuit was hidden in the sealed box, so as to teach infants to ask their caregiver to open it for them. The test phase (eight trials) was similar to the familiarization phase except for two elements. First, a variable memorization delay (3 or 12 s) was introduced. Second, the biscuits were now hidden half of the time in the unsealed box and the other half in the sealed box. Importantly, infants selected the sealed and unsealed boxes equally often ($t(21) = 0.3$; $p > 0.7$), showing that they could not discriminate the two boxes before pointing. Selection of the sealed box forced infants to either confirm their initial choice by asking their caregiver to open it or invalidate their initial choice by turning to the alternative box, whose content was directly reachable."*

Goupil & Kouider 2016

[151] *"In the current neuroscience literature there is a fair amount of confusion regarding how confidence is encoded in the brain. Some data indicate that confidence may be encoded by the same circuits involved in decision-making, others that confidence is monitored by dedicated structures...*

All the above studies suggest that confidence is implemented in regions that are not commonly considered as part of the decision-making circuitry, evoking the image of a looker inside the brain. Some of these studies, like Lau and Passingham (2006), Del Cul et al., (2009), Rounis et al., (2010) Fleming et al., (2010) McCurdy et al., (2013), Komura et al., (2013) and Lak et al., (2014) show a clear dissociation between performance and confidence, others like, Hebart et al., (2014) Middlebrooks and Sommer (2012) and Kepecs et al., (2008) show correlates of confidence in regions that are not traditionally considered to be involved in decision-making. Together these results suggest that there are separate and perhaps multiple areas involved in confidence monitoring and reporting. Future studies could be aimed at elucidating how these areas work together to form the circuit involved in monitoring and reporting confidence…

Although neurological, neuropsychological, fMRI and psychophysical data described above support the idea that confidence circuitry is separate from decision-making circuitry, recent electrophysiological experiments in monkeys suggest that these circuits are shared."

Grimaldi, Lau & Basso 2015

[152] *"Imagine that you are driving your car at night. There are no street lights on the road and your car's front lights are dim. As you are trying to keep the car on the road you need to determine which direction you and the other traffic are moving. This can be achieved by processing two distinct sensory inputs: the visual flow field created on the retina by your own motion and the vestibular stimulation, which measures acceleration. If the car in front of you suddenly brakes, you have to make a quick decision, based on these sensory inputs, about whether it is better to veer left or right (we are assuming that there are no additional obstacles or cars on either side, in which case the only important question is how to avoid a collision with the car ahead). The best decision requires determining whether your current heading is to the left or right side of the braking car and then to veer in that direction. The noise in the vestibular system as well as the glare of lights and random movement of cars creates uncertainty and, given these sources of stochasticity, you, or rather your brain, cannot know for sure the precise direction of heading. As this example illustrates, to perform well, the brain needs to be effective at dealing with a daunting array of uncertainties. Some originate in the external world, such as sensory or motor variability, whereas others are internal to the brain and are associated with cognitive variables, timing or abstract states. When dealing with these uncertainties, it is useful to represent current knowledge with probability distributions and update these on the basis of the rules of probabilistic inference—namely Bayes' theorem. Notably, there is ample experimental evidence that humans and other animals can indeed estimate and employ uncertainty to perform probabilistic inference about sensory, cognitive and motor variables (see ref. 9 for a review). In fact, in the particular case of heading direction, humans and animals have been shown to perform near optimally given the uncertainty inherent to the visual and vestibular information. There is also emerging evidence about how brains implement these uncertainty-based computations in neural circuits."*

Pouget et al. 2012

[153] *"The mathematical foundation of Bayesian concepts stems from the so-called Bayes' rule, named after one of its contributors, the 18th century British Reverend Thomas Bayes. Let's consider a practical example of how Bayes' rule works. A medical doctor faced with the following data D, a patient with a cough, contemplates three hypothetical diseases: a lung cancer (H_1), a cold (H_2) or gastroenteritis (H_3). The relative merit of each hypothesis can be deconstructed as follows according to Bayes' rule. Patients usually cough when afflicted by lung cancer or a cold but rarely in the case of gastroenteritis. Therefore, the likelihood of the potential cause for the cough is high under H_1 and H_2 and low under H_3. Second, a cold and gastroenteritis are much more prevalent diseases than lung cancer in the general population. The a priori likelihood of H_2 and H_3 is much higher than that of H_1. Given that only H_2 scores high both in a priori and current evidence, the most likely disease given the symptoms is a cold. Stated more generally, Bayes' rule says that our degree of belief in a hypothesis H given some current data D depends on the a priori likelihood of this hypothesis (what we know about it, independent of the current data), and the likelihood of the current data given this hypothesis. Formally, degrees of belief and likelihoods correspond to probabilities [1] and Bayes' rule reads:*

$p(H|D) = p(D|H)*p(H)/p(D)$.

Bayes' rule distinguishes between our belief a priori in the hypothesis p(H) and our belief in this hypothesis a posteriori, p(H|D), once particular data are considered to evaluate it. The notation p(D|H) is a shorthand for the probability of D given that we know H (the so-called likelihood of the data) and p(H|D) for the probability of H given that we know D."

Meyniel 2016

[154] *"A theoretical explanation of how confidence is encoded by the same neurons involved in decision-making is supported by the currently popular, Bayesian views of the brain (Friston, 2012; Lau, 2008; Pouget et al., 2013). Bayes theorem is a way to quantify uncertainty and is formally stated as: P(a/b) = P(b/a) P(b)/P(a) where P(a|b) is the conditional probability of event a occurring given the occurrence of event b, also called the posterior. P(b|a) is the conditional probability of observing event b given event a. This is also known as the likelihood. P(b) is the probability of event b also referred to as the prior. P(a) is a normalization term and for explanatory purposes can be ignored...*

Thinking about the brain in Bayesian terms is somewhat intuitive. Neurons, particularly those in sensory and motor areas have tuning curves; that is they can be described as radially symmetric functions of a stimulus parameter. Neurons show maximal discharge for optimal stimuli or movements and discharge less with stimulus or movement parameters that are less than optimal (Chalupa, 2003). Tuning curves are essentially, likelihood functions. They are measures of the probability of a particular outcome given a particular discharge rate (Foldiak, 1993; Jazayeri and Movshon, 2006; Sanger, 2002, 2003). Recent theoretical work (Ma et al., 2006), supported by the previously described experimental studies (Beck et al., 2008; Kim and Basso, 2010) shows that populations of neurons representing the likelihood and the prior, can be combined linearly in much the same way as Bayes' theorem combines two probability distributions, to provide a read-out of a decision in the form a posterior distribution…

A critical feature of this kind of an approach to understanding decision-making is that confidence or uncertainty is encoded implicitly across the population response or the posterior…

This kind of view, that perceptual decisions are computed in terms of Bayesian probability distributions in the brain, is one motivation for believing that confidence does not depend on specialized circuitry. If perceptual decisions are already computed in such probabilistic terms, confidence information should already be present in the circuits for decision-making… However, one concern is whether such information in the superior colliculus or LIP for example, can be read-out by those structures themselves. Even if the information is there, it is still possible that a monitoring module reads out the width of the distribution of the posterior."

Grimaldi, Lau & Basso 2015

"*A general understanding of the notion of confidence is that it fundamentally quantifies a degree of belief, or synonymously, a degree of reliability, trustworthiness, certitude, or plausibility. This common notion coincides closely with a formal one: that of Bayesian probability. Although a probability is sometimes considered to describe the likelihood of occurrence of random events in the world, from the viewpoint of an observer, whether such likelihoods constitute objective facts or reflect subjective knowledge is indistinguishable. Thus, probabilities simply are degrees of belief from the Bayesian viewpoint (Jaynes, 2003). Recognizing that much remains to be unpacked, we adopt the notion of Bayesian probability as the formal definition of subjective confidence."*

Meyniel, Sigman & Mainen 2015

"*A key claim of this review is that the notion of 'uncertainty' used in research on Bayesian neural computation (Fiser et al., 2010; Ma and Jazayeri, 2014; Pouget et al., 2013) and the notion of 'confidence' used in metacognitive research are two different manifestations of the same concept of Bayesian probability. First, we note that 'uncertainty' and 'confidence' are merely the inverse (or reciprocal) of one another, so the choice of emphasis is not an important difference. Instead, the critical difference is that 'confidence' in the metacognitive field is a single number, such as a numerical rating, whereas 'uncertainty' in the Bayesian computation field is a property of an array of numbers, such as a distribution of firing rates across neurons*."

Meyniel, Sigman & Mainen 2015

[155] *"We propose that confidence should be used to refer to the probability that a choice is correct, which we denote p(z = k|d = k, Image, Vestib). This definition has a long history in psychophysics and has been recently*

used in several studies. This is also what many authors call confidence, even if they don't always formally define it as such. This definition not only applies to decisions, but also to confidence in propositions, or potentially even to aspects of self-confidence. For example, suppose you are asked to express your confidence in the following proposition: 'Nigeria is the most populous African country.' This amounts to asking your confidence in choosing this proposition versus 'Nigeria is not the most populous African country.' Thus, as for decision confidence, the confidence in this proposition can be defined as the probability that the decision, 'Nigeria is the most populous African country', is correct. The same applies to some aspects of self-confidence. Lionel Messi is presumably highly self-confident in his ability to score in soccer games because the probability that the proposition 'I will score' (as opposed to 'I will not score') is correct tends to be high. The concept that unifies all of these seemingly different types of confidence is that they are about a choice being correct, even if only hypothetically, such that confidence can be expressed probabilistically by p(z = k|d = k, evidence). Here we focus mostly on confidence about decisions, but our conclusions apply just as well to propositions. When compared to the posterior p(z|Image, Vestib) over all possible choices, confidence is the probability mass of this posterior for one particular (overt or covert) choice. But does it ever make sense to maintain a separate measure of confidence rather than continuing to use the full posterior? In other words, why would you use a limited summary, confidence, when the entire posterior distribution is available? This is because confidence is in fact the only quantity that is needed in a wide variety of tasks. It is particularly important in sequential decisions, when subsequent choices depend on previous decisions. One example of such a task is a post-decision wager, in which subjects are asked to place a bet on whether their decision was correct. The optimal size of the wager, the investment, depends on the degree of belief that the initial choice was correct, with a higher wager when confidence is high. These types of post-decision wagers can be studied in the laboratory, even in animals. One example is a time investment task, initially introduced to study confidence in rats, that requires the decision maker to first gather evidence about which of several choice options is rewarded. After a choice is made, the reward is delayed for a randomized interval and it is up to the decision maker to choose how long to wait for this reward. To not wait in vain, it only makes sense to wait extended periods if the decision maker is confident of their choice. In fact, it can be shown that there is no need to store the posterior distribution over the choices for this kind of task: the probabilities associated with the choices that the subject did not select are irrelevant, the only required quantity is the probability that the selected choice is correct. Confidence can also be important for learning from feedback (Box 2) and group decision-making. However, confidence is not always the appropriate measure to use, even in sequential decisions. For instance, if a subject receives further information relevant to a previously taken choice, then the entire posterior distribution over the latent variable z, p(z|Image, Vestib), needs to be updated in the face of new evidence. Even in this situation, confidence may be a computationally efficient summary statistic to use instead of the full posterior distribution. Consider, for example, complex environments in which the posterior distribution might require an inordinate amount of data to learn, involve extremely complex inference to compute or require large neuronal resource to store. As a result, the posterior distributions computed will only be a rough approximation of the true posterior. Using confidence in these situations as an approximation to the full posterior can be the computationally appropriate strategy that beats other solutions that were optimal if more information were available."

Pouget et al. 2016

[156] "*Confidence has been extensively studied in multiple disciplines, primarily psychology and statistics, and the distinct conceptual frameworks of the two fields reveal confidence's dual manifestation as a subjective belief and an objective prediction. In psychology, confidence is often studied as a subjective feeling associated with beliefs about the world (Dunlosky and Metcalfe 2009; Kepecs and Mensh 2015). Introspection seems to be the only way to access this sense of confidence, and it is best communicated through verbalized self-reports. However, confidence is also a statistical quantity that can be defined as the likelihood that a belief is correct (Kepecs and Mainen 2012; Hangya et al. 2016; Pouget et al. 2016). This objective face of confidence measuring the degree of a belief can be precisely quantified using standard mathematical and statistical techniques. Indeed, confidence calculations lie at the foundation of statistical decision theory, machine learning, and hypothesis testing (Berger 1985).*"

Ott, Masset & Kepecs 2018

[157] "*… these theories have so far been explored and tested mostly in the domain of perceptual processing (Bejjanki et al., 2011, Berkes et al., 2011, Deneve et al., 1999, Fiser et al., 2010, Kim and Basso, 2010, Ma et al., 2006). It remains an open question to what extent probabilistic computation holds beyond low-level sensory and motor representations: e.g., the belief that 'it may rain tomorrow,' a reward expectation, etc. Forming probability distributions by simulating internal models could serve as the basis for a distributional neural*

representation of confidence in a variety of problems. There do exist a number of models for higher-level computations, for instance involving sampling schemes with integration of samples internally generated, e.g., for evaluating general-knowledge statements (Gigerenzer et al., 1991, Juslin et al., 2007, Koriat, 2012), for learning and goal-directed decisions (Hinton and Dayan, 1996, Legenstein and Maass, 2014, Solway and Botvinick, 2012), and even for probabilistic abstract reasoning (Chater et al., 2006, Denison et al., 2013, Vul et al., 2009)."

Meyniel, Sigman & Mainen 2015

"From this modest premise, our seemingly lofty aim is to bridge the gap between psychology on the one hand and neuroscience on the other. The foundation for our approach is first to recognize that, semantically, confidence is a property (degree, probability, etc.) that describes or modifies a referent (belief, response, memory, future event, etc.). Therefore it is impossible to refer precisely to confidence without specifying the object to which it pertains. In common usage the referent is often not made explicit and this is likely to contribute to conceptual confusion. We propose that the same general formal notion of confidence as Bayesian probability can be applied to widely different structures and processes. These include populations of neurons, neural functions, behavioral outputs, persons, etc."

Meyniel, Sigman & Mainen 2015

[158] *"But what is the relationship between the subjective sense of confidence in beliefs we experience as humans and the objective statistical computation that is also referred to as confidence? And because we are primarily interested in understanding the subjective sense of confidence, could we entirely avoid questions related to objective, statistical notions of confidence and simply focus on self-reported confidence measures? In other words, could we not just link human verbal reports of confidence to measures of neural activity and entirely sidestep issues related to statistical computations? Alas, confidence reports are often idiosyncratic: They are influenced by many factors, vary across contexts, and are inconsistent across individuals. For instance, the specific degrees of beliefs reported, whether low, high, or in-between levels of confidence, differs widely across individuals (Ais et al. 2016). The same decision situations can also lead to different confidence reports in one individual depending on context. At worst, self-reports do not provide useful predictions at all (Dunlosky and Metcalfe 2009). In these situations, identifying neural correlates of self-reported confidence would be expected to reveal a range of neural processes related to each individual's private notion of confidence and context. Therefore, although it is incumbent on any research on confidence to start from the human sense of confidence, there is a danger in taking confidence reports at face value, leading to a circular definition that makes it difficult to identify the neural underpinnings of confidence. But the challenge of linking subjective phenomenal experience with neural activity can be overcome by grounding self- reported confidence in objective statistical computations."*

Ott, Masset & Kepecs 2018

[159] *"To sum up, we have given different names (summary confidence, distributional confidence) to aspects of confidence that we think are worth keeping distinct. We have described how, for simple examples, summary confidence can be derived normatively from the distributional confidence information conveyed by probabilistic neural representations. We will go into more complexity later, with less direct routes and deviations from optimality (see A Brain-Scale, Hierarchical Neural Architecture for Confidence and also The Rough Edges). For the moment, the implications of this basic conceptualization can be related to the classic literature on confidence. We suggest that some confusion in the field of confidence studies is due to the conflation of distributional and summary forms. We propose that in decision-making, choice and confidence can be read out from the same neural representation (Kepecs and Mainen, 2012, Kepecs et al., 2008). This view resembles the "shared encoding" hypothesis reported by Grimaldi et al. (2015) or 'first-order model' (Timmermans et al., 2012) in which the same stream of information accounts for choice and confidence. However, these models are usually thought to entail that the same circuitry underpins choice and confidence (Grimaldi et al., 2015). We suggest the opposite: the mechanisms that read out a choice and a summary confidence from the same representation must be partly different, simply because they result in different things. Such 'parallel processing' of choice and confidence is the landmark of 'dual route models' (Timmermans et al., 2012), but our framework rejects a pure parallelism by assuming a common initial representation. Our view could therefore seem closer to 'hierarchical models' (Fleming and Dolan, 2012, Fleming et al., 2012, Timmermans et al., 2012). However, such models make a distinction between a first-order level (choice) and a second-order level (confidence) processing. This distinction is a landmark in the metacognition literature. In our view, there is no need for such a terminology: readout of choice and confidence are simply different without one being subordinate to the other."*

Meyniel, Sigman & Mainen 2015

[160] *"Interestingly, the readout of confidence can be selectively impaired in specific domains. Fleming and colleagues reported such a case: patients with brain lesions in the anterior PFC had preserved performance in the memory and perceptual domains and degraded confidence judgments specifically in the perceptual domain (Fleming et al., 2014). The fact that choice performance was preserved rules out the possibility that perception or memory, as a whole, were impaired, and points to the readout of confidence itself. This example suggests that one region alone does not suffice to read out confidence: at a minimum, it should involve a circuitry to collect specific inputs from different cognitive domains."*

Meyniel, Sigman & Mainen 2015

[161] *"The capacity to predict future events permits a creature to detect, model, and manipulate the causal structure of its interactions with its environment. Behavioral experiments suggest that learning is driven by changes in the expectations about future salient events such as rewards and punishments. Physiological work has recently complemented these studies by identifying dopaminergic neurons in the primate whose fluctuating output apparently signals changes or errors in the predictions of future salient and rewarding events. Taken together, these findings can be understood through quantitative theories of adaptive optimizing control."*

Schultz, Dayan & Montague 1997

[162] *"Dopamine neurons of the ventral tegmental area (VTA) and substantia nigra have long been identified with the processing of rewarding stimuli. These neurons send their axons to brain structures involved in motivation and goal-directed behavior, for example, the striatum, nucleus accumbens, and frontal cortex. Multiple lines of evidence support the idea that these neurons construct and distribute information about rewarding events…*

Surprisingly, after repeated pairings of visual and auditory cues followed by reward, dopamine neurons change the time of their phasic activation from just after the time of reward delivery to the time of cue onset. In one task, a naive monkey is required to touch a lever after the appearance of a small light. Before training and in the initial phases of training, most dopamine neurons show a short burst of impulses after reward delivery. After several days of training, the animal learns to reach for the lever as soon as the light is illuminated, and this behavioral change correlates with two remarkable changes in the dopamine neuron output: (i) the primary reward no longer elicits a phasic response; and (ii) the onset of the (predictive) light now causes a phasic activation in dopamine cell output. The changes in dopaminergic activity strongly resemble the transfer of an animal's appetitive behavioral reaction from the US to the CS. In trials where the reward is not delivered at the appropriate time after the onset of the light, dopamine neurons are depressed markedly below their basal firing rate exactly at the time that the reward should have occurred. This well-timed decrease in spike output shows that the expected time of reward delivery based on the occurrence of the light is also encoded in the fluctuations in dopaminergic activity (18)."

Schultz, Dayan & Montague 1997

"Dopamine neurons are therefore excellent feature detectors of the 'goodness' of environmental events relative to learned predictions about those events. They emit a positive signal (increased spike production) if an appetitive event is better than predicted, no signal (no change in spike production) if an appetitive event occurs as predicted, and a negative signal (decreased spike production) if an appetitive event is worse than predicted."

Schultz, Dayan, Montague 1997

[163] *"What has been learned must sometimes be unlearned in a changing world. Yet knowledge updating is difficult since our world is also inherently uncertain. For instance, a heatwave in winter is surprising and ambiguous: does it denote an infrequent fluctuation in normal weather or a profound change? Should I trust my current knowledge, or revise it? We propose that humans possess an accurate sense of confidence that allows them to evaluate the reliability of their knowledge, and use this information to strike the balance between prior knowledge and current evidence. Our functional MRI data suggest that a frontoparietal network implements this confidence-weighted learning algorithm, acting as a statistician that uses probabilistic information to estimate a hierarchical model of the world*."

Meyniel & Dehaene 2017

[164] *"Metacognition, the ability to internally evaluate our cognitive processes, is critical for adaptive behavioral control, particularly as many real-life decisions lack immediate feedback. Specifically, action outcomes can be ambiguous, delayed, occur only after a sequence of subsequent decisions, or might never occur at all. Yet*

behavioral and neural evidence indicate that subjects are able to evaluate their choices online in the absence of immediate feedback, forming estimates of decision confidence and detecting and correcting response errors."

Rouault, Dayan & Fleming 2019

[165] *"Although both JOLs and metacomprehension judgments typically demonstrate above-chance relative accuracy, the calibration and relative accuracy of these judgments is not always impressive (Maki 1998b, Nelson & Dunlosky, 1991; Weaver & Kelemen, 1997)."*

Serra & Metcalfe 2009

[166] *"These results show that confidence and perceptual performance can be dissociated and indicate that decision and confidence are generated by different specific rules, and therefore may have different neural substrates."*

Grimaldi, Lau & Basso 2015

"If neural circuits function inherently probabilistically, why is confidence sometimes estimated in a way that is inconsistent with probability theory, reflecting biases and reliable inconsistencies (Kahneman, 2013)? Many behaviors are close to optimal (Ma and Jazayeri, 2014; Maloney and Zhang, 2010; Pouget et al., 2013) but decades of experimentation in 'real-life' decision problems have also shown that humans commonly assign confidence sub-optimally, relying on sub-samples of the data, focusing on tokens (representative exemplars), ignoring the variance of the distribution, and over-weighting evidence confirming commitments and choices that have been made (Griffin and Tversky, 1992; Kahneman, 2013). One simple possibility is that errors of confidence are found in real life because real life involves complex high dimensional problems for which there may not be accurate distributional representations available in neural circuits. However, distortions of confidence can be observed even in simple decision tasks (Graziano and Sigman, 2009; Jarvstad et al., 2013; Rahnev et al., 2012; Wu et al., 2009; Zylberberg et al., 2012, 2014). Distortions of confidence were mentioned already in the seminal work of Peirce and Jastrow on small differences in tactile perception (Peirce and Jastrow, 1884). Here we argue that suboptimality arises from approximations inherent in probabilistic representations, particularly in the readout of summary confidence from probabilistic representations, and that this framework can potentially help to explain the specific deviations from optimality that arise as a consequence of specific features of these approximations…

Deviations from optimality observed in behavior are often used as reasons to reject a probabilistic view of brain functioning. The examples above show that decision making, even in simple perceptual problems, can follow a probabilistic logic and still be suboptimal because of specific approximations (Griffiths et al., 2012; Ma and Jazayeri, 2014). Acerbi and colleagues showed in a sensory motor task that prior distributions with different shapes (Gaussian, non-symmetric, bimodal, etc.) could be used with only minor errors that, crucially, were independent from the shape of the distribution (Acerbi et al., 2014). Their results suggest that deviations from optimality observed in the behavior are not due to a fundamental inability to represent and combine probability distributions, but might instead be due to random noise in this process…

In summary, what we perceive as heuristics in confidence judgments may result from different sources: (1) genuine approximations of a read-out process, including issues of calibration; (2) applying stereotyped read-out procedures that make certain assumptions that do not hold in a given context, a form of approximation referred as relaxation; and (3) using variables that covary with the relevant neural confidence information in cases in which this information is not accessible for explicit reports."

Meyniel, Sigman & Mainen 2015

[167] *"Much research on metacognition has focused on the accuracy of monitoring judgments, which has mainly been conceptualized in two ways: calibration and relative accuracy. The calibration of one's judgments refers to a difference score between the mean of one's predictive judgments and one's performance on the task being judged or predicted. Consider hypothetical participants in a laboratory study on metamemory (metacognition about memory). While studying paired associate items (e.g., two unrelated nouns) for the criterion test, the participants make JOLs on a 0%-100% scale indicating the percent-likelihood that they will correctly remember each item on a test. Suppose that the overall mean of their JOLs is 88%. Assuming that the participants correctly remember a mean of 66% of the items on the test, their calibration score will be +22%, indicating overconfidence. A group of participants in such a study would be said to demonstrate 'good' calibration if the overall mean of their JOLs did not significantly differ from the overall mean of their performance scores. The relative accuracy of one's judgments refers to a measure of how well one's judgments differentiate performance*

on the cognitive task being judged. This measure is usually calculated by computing a gamma correlation between one's judgments about individual items and performance on those same test items. Like a Pearson correlation, gamma can range from –1.0 to +1.0. If a hypothetical participant gives mostly high JOLs to items they will remember on the test and low JOLs to items they will not remember on the test, their gamma correlation will be positive. Doing the opposite will result in a negative correlation. Assigning JOLs to items at random tends to result in a correlation of zero (or a correlation might even be incalculable). Typically, a gamma correlation is calculated for each participant. The mean gamma correlation is then calculated across all of the participants in a group to estimate the relative accuracy of their judgments; if the mean is significantly greater than zero, the group's judgments are said to be above chance."

Serra & Metcalfe 2009

[168] *"To study the 'meaningfulness' of confidence ratings, researchers often assess the across-trial correlation between accuracy and confidence; a high degree of correlation is interpreted as an indication of a high degree of reliability of the measurement of confidence (Kornell et al., 2007), or in short, high metacognitive sensitivity.*

However, the raw statistical correlation between accuracy and confidence is not necessarily a precise measure of metacognitive sensitivity. The correlation is a function of both how well the subject can distinguish between correct and incorrect trials, as well as the overall propensity to give high or low confidence responses. If subjects are very liberal, reporting high confidence frequently, or if they are very conservative, often reporting low confidence, then the interpretation of the correlation strength is limited. As it happens for quantifying perceptual judgment, likewise, the application of signal detection theory (SDT) in this case provides a more reliable approach to quantifying metacognitive sensitivity (Galvin et al., 2003; Maniscalco and Lau, 2012)."

Grimaldi, Lau & Basso 2015

"*Early cognitive psychologists were interested in how well people could assess or monitor their own knowledge, and asking for confidence ratings was one of the mainstays of psychophysical analysis (Peirce and Jastrow, 1885). For example, Henmon (1911) summarized his results as follows: 'While there is a positive correlation on the whole between degree of confidence and accuracy the degree of confidence is not a reliable index of accuracy.' This statement is largely supported by more recent research in the field of metacognition in a variety of domains from memory to perception and decision-making: subjects have some metacognitive sensitivity, but it is often subject to error (Nelson and Narens, 1990; Metcalfe and Shimamura, 1996)…*

From the outset, it is important to distinguish two aspects, namely sensitivity and bias. Metacognitive sensitivity is also known as metacognitive accuracy, type 2 sensitivity, discrimination, reliability, or the confidence-accuracy correlation. Metacognitive bias is also known as type 2 bias, over- or underconfidence or calibration.

…metacognitive sensitivity is the ability to discriminate correct from incorrect judgments…

Metacognitive bias is the tendency to give high confidence ratings, all else being equal. The simplest of such measures is the percentage of high confidence trials ..., or the average confidence rating over trials…

A useful starting point for all the measures of metacognitive sensitivity that follow is the 2 × 2 confidence-accuracy table (Table 1). This table simply counts the number of high confidence ratings assigned to correct and incorrect judgments, and similarly for low confidence ratings. Intuitively, above-chance metacognitive sensitivity is found when correct trials are endorsed with high confidence to a greater degree than incorrect trials…

The simplest measure of association between the rows and columns of Table 1 is the phi (ϕ) correlation. In essence, phi is the standard Pearson r correlation between accuracy and confidence over trials. That is, if we code correct responses as 1's, and incorrect responses as 0's, accuracy over trials forms a vector, e.g., [0 1 1 0 0 1]. And if we code high confidence as 1, and low confidence as 0, we can likewise form a vector of the same length (number of trials). The Pearson r correlation between these two vectors defines the "phi" coefficient. A related and very common measure of metacognitive sensitivity, at least in the memory literature, is the Goodman–Kruskall gamma coefficient, G (Goodman and Kruskal, 1954; Nelson, 1984). In a classic paper, Nelson (1984) advocated G as a measure of metacognitive sensitivity that does not make the distributional assumptions of SDT.

G can be easily expanded to handle designs in which confidence is made using a rating scale rather than a dichotomous high/low design (Gonzalez and Nelson, 1996). Though popular, as measures of metacognitive sensitivity both phi and gamma correlations have a number of problems. The most prominent is the fact that both can be 'contaminated' by metacognitive bias. That is, for subjects with a high or low tendency to give high confidence ratings overall, their phi correlation will be altered (Nelson, 1984)[2]. Intuitively one can consider the

extreme cases where subjects perform a task near threshold (i.e., between ceiling and chance performance), but rate every trial as low confidence, not because of a lack of ability to introspect, but because of an overly shy or humble personality. In such a case, the correspondence between confidence and accuracy is constrained by bias. In an extensive simulation study, Masson and Rotello (2009)showed that G was similarly sensitive to the tendency to use higher or lower confidence ratings (bias), and that this may lead to erroneous conclusions, such as interpreting a difference in G between groups as reflecting a true underlying difference in metacognitive sensitivity despite possible differences in bias."

Fleming & Lau 2014

[169] *"We are constantly flooded with information that helps form our beliefs about reality (e.g., via the Web, advertising, colleagues, and friends). Understanding how beliefs are formed is critical, as beliefs drive our actions and decisions. Normative theories assume beliefs are adjusted according to Bayes' Rule. Indeed, this assumption often holds when people in- corporate favorable news into their existing beliefs. However, for unfavorable news people show an aversion to incorporating new information. Specifically, they discount the strength of the new information leading to noisy posterior beliefs. This tendency to selectively ignore negative information is known as the 'good news/bad news effect'. For example, people adjust their beliefs regarding their level of intelligence and physical attractiveness when they receive in- formation indicating they are more intelligent and attractive than they had assumed. However, they relatively fail to adjust their beliefs in response to information suggesting they rate lower on these attributes than they had previously thought. In addition, when learning that their risk of experiencing future negative events, such as cancer, is higher than they had expected, people are less likely to integrate these data into their prior beliefs relative to a situation when they learn that their risk is lower than expected."*

Sharot et al. 2012

[170] *"We explored these predictions in four studies. In each, we presented participants with tests that assessed their ability in a domain in which knowledge, wisdom, or savvy was crucial: humor (Study 1), logical reasoning (Studies 2- and 4), and English grammar (Study 3). We then asked participants to assess their ability and test performance. In all studies, we predicted that participants in general would overestimate their ability and performance relative to objective criteria. But more to the point, we predicted that those who proved to be incompetent (i.e., those who scored in the bottom quarter of the distribution) would be unaware that they had performed poorly. For example, their score would fall in the 10th or 15th percentile among their peers, but they would estimate that it fell much higher (Prediction 1). Of course, this overestimation could be taken as a mathematical verity. If one has a low score, one has a better chance of overestimating one's performance than underestimating it. Thus, the real question in these studies is how much those who scored poorly would be miscalibrated with respect to their performance. … Indeed, across the four studies, participants in the bottom quartile not only overestimated themselves, but thought they were above-average … What causes this gross overestimation? Studies 3 and 4 pointed to a lack of metacognitive skills among less skilled participants. Bottom-quartile participants were less successful than were top- quartile participants in the metacognitive tasks of discerning what one has answered correctly versus incorrectly (Study 4) and distinguishing superior from inferior performances on the part of one's peers (Study 3). More conclusively, Study 4 showed that improving participants' metacognitive skills also improved the accuracy of their self-appraisals. Note that these findings are inconsistent with a simple regression effect interpretation of our results, which does not predict any changes in self-appraisals given different levels of metacognitive skill. Regression also cannot explain the fact that bottom-quartile participants were nearly 4 times more miscalibrated than their top-quartile counterparts.*

Study 4 also revealed a paradox. It suggested that one way to make people recognize their incompetence is to make them competent. Once we taught bottom-quartile participants how to solve Wason selection tasks correctly, they also gained the metacognitive skills to recognize the previous error of their ways. Of course, and herein lies the paradox, once they gained the metacognitive skills to recognize their own incompetence, they were no longer incompetent."

Kruger & Dunning 1999

[171] *"Kruger and Dunning (1999) suggested that, across many intellectual and social domains, it is the poorest performers who hold the least accurate assessments of their skill and performances, grossly overestimating how well their performances stack up against those of their peers. For example, students performing in the bottom 25% among their peers on tests of grammar, logical reasoning, and humor tended to think that they are performing above the 60th percentile (Kruger & Dunning, 1999). Further, this pattern has been conceptually replicated among undergraduates completing a classroom exam (Dunning, Johnson, Ehrlinger, & Kruger,*

2003), medical students assessing their interviewing skills (Hodges, Regehr, & Martin, 2001) clerks evaluating their performance (Edwards, Kellner, Sistrom, & Magyari, 2003), and medical lab technicians evaluating their on-the-job expertise (Haun, Zeringue, Leach, & Foley, 2000)."

Ehrlinger et al. 2008

[172] *"In essence, we argue that the skills that engender competence in a particular domain are often the very same skills necessary to evaluate competence in that domain—one's own or anyone else's. Because of this, incompetent individuals lack what cognitive psychologists variously term metacognition (Everson & Tobias, 1998), metamemory (Klin, Guizman, & Levine, 1997), metacomprehension (Maki, Jonas, & Kallod, 1994), or self-monitorings kills (Chi, Glaser, & Rees, 1982). These terms refer to the ability to know how well one is performing, when one is likely to be accurate in judgment, and when one is likely to be in error. For example, consider the ability to write grammatical English. The skills that enable one to construct a grammatical sentence are the same skills necessary to recognize a grammatical sentence, and thus are the same skills necessary to determine if a grammatical mistake has been made. In short, the same knowledge that underlies the ability to produce correct judgment is also the knowledge."*

Kruger & Dunning 1999

[173] *"Several lines of research are consistent with the notion that incompetent individuals lack the metacognitive skills necessary for accurate self-assessment. Work on the nature of expertise, for instance, has revealed that novices possess poorer metacognitive skills than do experts. In physics, novices are less accurate than experts in judging the difficulty of physics problems (Chi et al., 1982). In chess, novices are less calibrated than experts about how many times they need to see a given chessboard position before they are able to reproduce it correctly (Chi, 1978). In tennis, novices are less likely than experts to successfully gauge whether specific play attempts were successful (McPherson & Thomas, 1989). These findings suggest that unaccomplished individuals do not possess the degree of metacognitive skills necessary for accurate self-assessment that their more accomplished counterparts possess. However, none of this research has examined whether metacognitive deficiencies translate into inflated self-assessments or whether the relatively incompetent (novices) are systematically more miscalibrated about their ability than are the competent (experts)."*

Kruger & Dunning 1999

[174] *"We report two studies in which participants completed a popular performance-based measure of analytic-thinking disposition, the CRT (Frederick, 2005), and were subsequently asked to estimate how many of the items they had gotten correct (Mata, Ferreira, & Sherman, 2013; Noori, 2016). Following Kruger and Dunning (1999), we hypothesized that participants who performed poorly on the CRT would overestimate their performance to a greater extent than would those who performed well (i.e., less-analytic people should be more poorly calibrated). In addition, participants were also asked to self-report their need or desire to think analytically using the NC scale. We predicted a Dunning–Kruger effect, such that participants who performed particularly poorly on the CRT (indicating an intuitive or non-analytic-thinking disposition) would overreport the degree to which they were disposed to analytic thinking. In Study 2, we used an independent assessment of analytic thinking—the heuristics-and-biases inventory (Toplak, West, & Stanovich, 2011, 2014)—to assess whether nonanalytic individuals are genuinely worse at recognizing their bias. Put differently, participants who were decidedly nonanalytic based on the performance measure should be less-suited to assess their degree of analyticity on a self-report measure, leading to poor calibration in terms of both estimated CRT accuracy and self-reported NC. Our results provide empirical support for Dunning–Kruger effects in both estimates of reasoning performance and self- reported thinking disposition. Particularly intuitive individuals greatly overestimated their performance on the CRT—a tendency that diminished and eventually reversed among increasingly analytic individuals. Moreover, self-reported analytic- thinking disposition—as measured by the Ability and Engagement subscales of the NC scale—was just as strongly (if not more strongly) correlated with estimated CRT performance than with actual CRT performance. In addition, an analysis using an additional performance-based measure of analytic thinking—the heuristics-and-biases battery—revealed a systematic miscalibration of self-reported NC, where- in relatively intuitive individuals report that they are more analytic than is justified by their objective performance. Together, these findings indicate that participants who are low in analytic thinking (so-called intuitive thinkers) are at least somewhat unaware of (or unresponsive to) their propensity to rely on intuition in lieu of analytic thought during decision making. This conclusion is consistent with previous research that has suggested that the propensity to think analytically facilitates metacognitive monitoring during reasoning (Pennycook et al., 2015b; Thompson & Johnson, 2014). Those who are genuinely*

analytic are aware of the strengths and weaknesses of their reasoning, whereas those who are genuinely nonanalytic are perhaps best described as 'happy fools' (De Neys et al., 2013)."

Pennycock et al. 2017

[175] *"The present study focuses on this critical error sensitivity issue. Preschool children were given a classic version of a number conservation task in which an intuitively cued response conflicted with the correct conservation response and a control version in which this conflict was not present. After solving each version children were asked to indicate their response confidence. Results showed that in contrast with children who gave a correct conservation response, preschoolers who erred showed a sharp confidence decrease after solving the classic conflict problem. This suggests that non-conserving preschoolers detect that their response is questionable and are less ignorant about conservation than their well-documented errors might have previously suggested."*

De Neys, Lubin & Houdé 2014

[176] *"This contraposition indicates that the neurological (e.g., De Neys et al., 2008) and physiological (De Neys et al., 2010) conflict detection signals may be relatively effective, but the response to this signal may actually be rather ineffective. De Neys et al. (2013) found a large (15%) decrease in confidence for the bat-and-ball problem relative to a control, but 82% confidence is still quite high."*

Pennycock et al. 2017

[177] *"Although it is clear that people are often biased, the nature of this bias is poorly understood. Some authors claim that people reason heuristically by default and that most of the time they are simply not aware that their intuitions might be wrong. The dominance of intuitive thinking is attributed to a failure to monitor the output of the heuristic reasoning process. In this view, because of lax monitoring, people fail to detect that an intuitive response conflicts with the response favored by probability. The problem is that people do not know that their judgment is biased. This view has been popularized by the work of authors such as Kahneman (2002) and Evans (1984, 2003)."*

De Neys, Vartanian & Goel (2008)

[178] *"Intuitive or lay theories are thought to influence almost every facet of everyday cognition. People appeal to explanatory relations to guide their inferences in categorization, diagnosis, induction, and many other cognitive tasks, and across such diverse areas as biology, physical mechanics, and psychology (Gopnik & Wellman, 1994; Keil, 1998; Murphy & Medin, 1985; Murphy, 2000). Individuals will, for example, discount high correlations that do not conform to an intuitive causal model but overemphasize weak correlations that do (Chapman & Chapman, 1969). Theories seem to tell us what features to emphasize in learning new concepts as well as highlighting the relevant dimensions of similarity (Murphy, 2002). Intuitive theories have also been heavily emphasized in accounts of the cognitive development of children (Gelman & Koenig, 2002) and even of infants (Spelke, Breinliinger, Macomber, & Jacobson, 1992). Concepts seem to be embedded within larger sets of explanatory relations that are essential to understanding the structure of the concepts themselves, how they are learned, and how they change over time. But even as theories have become more central to the study of concepts, it is also now evident that folk theories are rarely complete or exhaustive explanations in a domain (Wilson & Keil, 1998). Indeed, even the theories used daily to guide scientific research are now considered to be incomplete, or at least less formally logical than classical views assumed them to be (Boyd, 1991; Salmon, 1989, 1998). Science-in-practice is often driven by hunches and vague impressions."*

Rozenblit & Keil 2002

[179] *"The pattern of results so far indicates a special difficulty in calibrating one's explanatory knowledge about devices. In the studies with procedures and movies, the participants were well calibrated. In the study with factual knowledge about capitals, the participants were overconfident but markedly less so than with explanatory knowledge of devices… To summarize, studies with devices and natural phenomena both show large drops in knowledge estimates. Procedures and Narratives show no drop, while Geography Facts show only a small drop. The results demonstrate large differences in knowledge calibration across knowledge domains, casting serious doubt on the meaningfulness of 'general overconfidence' about knowledge. The studies also raise intriguing possibilities about the mechanism behind over-confidence, which we address in the next few studies…One conclusion that can be drawn from this research is that the well-established blanket approach to overconfidence with 'general knowledge' is almost certainly misleading. Large inter-domain differences in calibration imply that structural properties of knowledge have a powerful impact on the process of knowledge*

assessment. 'General knowledge' is a chimera— a mythological composite creature. Taking it seriously distracts from interesting questions about how knowledge assessment works, and the theoretically important issues of how the structural properties of knowledge influence calibration."

Rozenblit & Keil 2002

[180] *"To clarify the distinctive nature of our proposal it is useful to briefly consider prior research on overconfidence. Relevant research in the judgment and decision-making tradition has used the disparity between people's average confidence levels for their answers to almanac questions and the proportion of correct answers to argue that people are overconfident (Fischhoff, 1982; Lichtenstein & Fischhoff, 1977; Yates, Lee, & Shinotsuka, 1996; Yates, Lee, & Bush, 1997). This tradition, however, does not focus on how illusions of knowing might differ across kinds of knowledge. Lumping diverse kinds of knowledge into a hypothetical 'general knowledge' and looking for an overall overconfidence effects may well obscure large differences in calibration between knowledge types. The cognitive psychology literature on text comprehension also suggests overconfidence about one's knowledge. People are often poor at detecting when they have failed to understand a piece of text, both as adults (Glenberg & Epstein, 1985; Glenberg, Wilkinson, & Epstein, 1982; Lin & Zabrucky, 1998) and as children (Markman, 1977; Markman, 1979). In contrast, the current studies are concerned with people's ability to assess the knowledge they have before coming into the lab, rather than things learned in the course of an experiment. The implications of our research are different: they tell us less about how people learn when reading, and more about individuals' intuitive theories about how knowledge is stored and about the mismatch between what they think they already know and what they really know. Another area of research has focused on meta-cognition and feelings of knowing (FOK) (Koriat, 1995; Metcalfe, Schwartz, & Joaquim, 1993). One recent analysis considers the two main models for FOK to be the cue familiarity model and the accessibility model (Koriat & Levy-Sadot, 2001). The accessibility model claims that the ease of accessing information prompted by the target drives FOKs. The cue familiarity model claims that FOK judgments are elicited by the familiarity of the cues themselves…*

Overconfidence also exists in areas that have little to do with knowledge. Participants have been shown to be overconfident about their future performance on motor tasks (e.g., West & Stanovich, 1997), their abilities compared to other people's abilities (e.g., Kruger & Dunning, 1999), and about their competence to perform a broad range of tasks (Bjork, 1998)."

Rozenblit & Keil 2002

"Folk theories, we claim, are even more fragmentary and skeletal, but laypeople, unlike some scientists, usually remain unaware of the incompleteness of their theories (Ahn & Kalish, 2000; Dunbar, 1995; diSessa, 1983). Laypeople rarely have to offer full explanations for most of the phenomena that they think they understand. Unlike many teachers, writers, and other professional 'explainers,' laypeople rarely have cause to doubt their naive intuitions. They believe that they can explain the world they live in fairly well. They are novices in two respects. First, they are novice 'scientists'—their knowledge of most phenomena is not very deep. Second, they are novice epistemologists—their sense of the properties of knowledge itself (including how it is stored) is poor and potentially misleading.… We argue here that people's limited knowledge and their misleading intuitive epistemology combine to create an illusion of explanatory depth (IOED). Most people feel they understand the world with far greater detail, coherence, and depth than they really do. The illusion for explanatory knowledge–knowledge that involves complex causal patterns—is separate from, and additive with, people's general overconfidence about their knowledge and skills. We therefore propose that knowledge of complex causal relations is particularly susceptible to illusions of understanding."

Rozenblit & Keil 2002

[181] *"People's estimations of the future are often unrealistically optimistic. A problem that has puzzled scientists for decades is why human optimism is so pervasive when reality continuously confronts us with information that challenges these biased beliefs. According to influential learning theories, agents should adjust their expectations when faced with disconfirming information. However, this normative account is challenged by observations that providing people with evidence that disconfirms their positive outlook often fails to engender realistic expectations. For example, highlighting previously unknown risk factors for diseases is surprisingly ineffective at altering an individual's optimistic perception of their medical vulnerability. Even experts show worrying optimistic biases. For instance, financial analysts expect improbably high profits and family law attorneys underestimate the negative consequences of divorce. The wider societal importance of these errors derives from the fact that they reduce precautionary actions, such as practicing safe sex or saving for retirement, On the upside, optimistic expectations can lower stress and anxiety, thereby promoting health and well-being.*

Although the existence of unrealistic optimism has been extensively documented the biological and computational principles that help to maintain optimistically biased predictions in the face of reality are unknown. Notably, such biases are not explained by theories assuming equal learning across outcome valence."

Sharot, Korn & Dolan 2011

[182] *« Que se passe-t-il dans le cas particulier où l'apprenant n'a pas d'expérience acquise, comme c'est le cas des jeunes enfants ? Dans ce cas, ils se montrent sur-confiants, c'est-à-dire qu'ils ont peu ou pas de doutes sur le fait qu'ils vont réussir leur action, et ne s'attendent pas non plus à ce que des efforts importants et soutenus soient nécessaires. Il est généralement admis que la sur-confiance systématique des grands débutants a une fonction adaptative, leur permettant de s'engager plus volontiers dans des tâches cognitives nouvelles. Comme la sur-confiance n'est pas confirmée par les résultats, elle est suivie par une phase de sous-confiance, qui met en péril la motivation à persister dans l'apprentissage.* »

Proust 2019

[183] *"In many situations, the low predictability of these judgments can be attributed to participants' basing them on information that is not diagnostic of their future performance. As described by Metcalfe, Schwartz, and Joaquin (1993):*

> *[When] making judgments related to the external world rather than to their internal states and abilities, people use other heuristics or rules of thumb. These heuristics approximate the uncertain quantity indirectly, rather than measuring the quantity itself. Because they do not measure the quantity directly, such heuristics may result in biases and errors. (p. 860).*

In other situations (e.g., delayed JOLs, as we discuss later), calibration and relative accuracy can be quite good because participants base their judgments on information that is diagnostic of their future performance. This section reviews some heuristics that are sometimes used to inform metacognitive judgments and that can produce errors and illusions in monitoring. This section is not meant to describe all of the potential errors and poor heuristics that might arise in metacognitive monitoring, but simply to provide some illustrative examples. It also suggests some ways that the errors associated with these particular cues might materialize in the classroom. As we learn more about heuristics that lead to accurate and inaccurate metacognitive judgments, we can better understand which heuristics aid or hinder metacognitive accuracy."

Serra & Metcalfe 2009

[184] *"Learners' familiarity with the information being judged can have an influence on their metacognitive judgments. Because this experience of familiarity often results from prior exposures to and learning of the information, it can be diagnostic of a greater likelihood that the information is known (i.e., it often results in accurate judgments). Unfortunately, familiarity can also arise in situations where it is unrepresentative of knowledge. For example, participants in a study by Reder and Ritter (1992) solved difficult arithmetic problems. After each was presented, they had to quickly choose whether to calculate the answer to each problem or recall it from memory (calculation, of course, was the only option the first time a problem was presented). Participants received 50 points for correctly recalling an answer and 5 points for correctly calculating an answer (but only if the selections and responses were made within the designated time limits). Participants were later paid .05 cents for each point earned. By manipulating the occurrence of specific numbers in the set of problems, Reder and Ritter manipulated the participants' familiarity for the numbers present in the problems independent of their memory for specific problems and their solutions. Participants were able to use their familiarity with the numbers—both independently and as whole problems—to quickly decide whether they knew the answers to the problems. This strategy proved helpful when the specific problem had actually been presented in accordance with the participants' familiarity for the numbers, but proved to be faulty when familiar numbers were combined into novel problems—problems for which the participants could not actually recall an answer."*

Serra & Metcalfe 2009

[185] *"Several studies have demonstrated that judgments of learning are sometimes made based on retrieval fluency—the speed and probability with which information is brought to mind (Benjamin, Bjork, & Schwartz, 1998). Retrieval fluency is predictive of memory performance in many situations (e.g., Serra & Dunlosky, 2005), so using it to inform metacognitive judgments will often be appropriate. It has recently been suggested that the utility of cues such as retrieval fluency and ease of learning are automatically incorporated into metacognitive judgments based on their ecological validity at the time the judgment is being made (Koriat, 2008). For example, Koriat (2008) recently described the easily learned, easily remembered (ELER) heuristic, which stems from the*

observation that items that are easily learned are typically remembered better than items that are difficult to learn. In the studies reported by Koriat (2008), participants apparently used ease of learning as a cue to inform their metacognitive judgments in a way that was directly correlated with the validity of that cue."

Serra & Metcalfe 2009

[186] *"When participants are required to list the reasons for their choice of an answer to almanac questions (Koriat et al., 1980, 2008), they typically mention logical or rational considerations. However, it is clear that some of the representations that tip the balance in favor of one choice or the other consist of associations and images that cannot be expressed in a propositional form, and some operate bellow full consciousness. Indeed, studies of the illusory-truth effect indicate that the mere familiarity and fluency of a statement that are caused by its repetition or by its context can influence the perceived truth of that statement (Arkes, Hackett, & Boehm, 1989; Bacon, 1979; Hasher, Goldstein, & Toppino, 1977; Unkelbach & Stahl, 2009)."*

Koriat 2012

[187] *"The overconfidence bias is assumed to follow from the basic assumption that confidence judgments rely on reliability as a cue for validity. Reliance on reliability—the consistency with which a choice is supported—may instill inflated confidence because reliability is practically always higher than validity. Indeed, although confidence judgments yielded an overconfidence bias when evaluated against correctness, these judgments were not markedly inflated when evaluated against several indexes of self- consistency. It should be stressed that the SCM account of overconfidence does not postulate a specific bias like the biases pro- posed in previous discussions (Koriat et al., 1980; Ronis & Yates, 1987; Taylor & Brown, 1988). Rather, the bias is assumed to stem from the very basis of confidence judgments."*

Koriat 2012

[188] *"Current-knowledge heuristic. Sometimes, after information has been obtained or understood, people think they knew or understood it all along. Participants in a study by Fischhoff (1975) read passages detailing outcomes associated with historical events (such as a battle). They then judged if they would have predicted the outcome before reading the passage. Fischhoff demonstrated that these participants could not avoid using their knowledge of the outcome when making this judgment; they even judged that they would have correctly predicted highly unlikely events (some of which were false). The participants in his study demonstrated hindsight bias—a tendency to use new knowledge when thinking about the past—without even knowing that they were doing so."*

"Association heuristics. Some information seems easier to understand or remember when it is studied than it will actually be to remember or apply later on a test. Koriat and Bjork (2006) termed such an illusion foresight bias and demonstrated a type of paired-associate that produces such an effect. These pairs were composed of two words that had a strong backwards association but a weak forward association (i.e., one would be likely to think of the first word when shown the second word but not likely to think of the second word when shown the first word). For example, consider the pair 'fire—blaze'. The word 'blaze' is almost always freely-associated to the word 'fire', but 'fire' rarely—if ever—is freely-associated to the word 'blaze'. When such pairs are studied and judged in a typical metamemory procedure, the presence of both words at study makes them seem highly related. At test, however, the stimulus word (fire) is not actually likely to produce the response word (blaze). The association strength present at study produces the illusion that the response word will easily be recalled at test."

"Heuristics that can cause illusions of knowing. One's experience with learning materials sometimes causes the illusion that the materials have been understood when in fact they have not. Participants in a study by Glenberg, Wilkinson, and Epstein (1982) demonstrated an illusion of knowing (i.e., their judgments were overconfident) when asked to rate their comprehension for texts containing factual contradictions. Participants often failed to find these contradictions yet rated their understanding of the texts as being high. This even occurred when factual contradictions were in two adjacent sentences (Glenberg et al., 1982). These findings suggest that readers do not attempt to monitor their understanding across a whole text, but rather at lower levels such as at the per-sentence level."

Serra & Metcalfe 2009

[189] *"As described in the previous section of this chapter, metacognitive judgments are prone to errors, biases, and metacognitive illusions. Experience with and information about these illusions can help to reduce some of these biases. Koriat and Bjork (2006) described one such illusion—foresight bias—in which some to-be-studied pairs of words have a strong backwards association but a weak forward association (e.g., fire—blaze). The*

presence of both words at study but not at test produced overconfident JOLs for these items. As in King, Zechmeister, and Shaughnessy (1980), Koriat and Bjork (2006) demonstrated that study-test practice reduced this bias, but it did not transfer to new items. Explicit training about the foresight bias and the type of item that causes it, however, not only reduced the bias but also transferred to new items."

Serra & Metcalfe 2009

[190] *"Most theories predict that when people indicate that they are highly confident they are producing their strongest responses. Hence, if such a high confidence response is in error it should be overwritten only with great difficulty. In contrast to this prediction, we have found that people easily correct erroneous responses to general information questions endorsed as correct with high- confidence, so long as the correct answer is given as feedback. Three potential explanations for this unexpected hypercorrection effect are summarized. The explanation that is tested here, in two experiments, is that after a person commits a high-confidence error the correct answer feedback, being surprising or unexpected, is given more attention than is accorded to the feedback to low-confidence errors. This enhanced attentional capture leads to better memory. In both experiments, a tone detection task was presented concurrently with the corrective feedback to assess the attentional capture of feedback stimuli. In both, tone detection was selectively impaired during the feedback to high confidence errors. It was also negatively related to final performance, indicating that the attention not devoted to the tone detection was effectively engaged by the corrective feedback. These data support the attentional explanation of the high-confidence hypercorrection effect."*

Butterfield & Metcalfe 2006

[191] *"The purpose of the current research is to reexamine the effect of feedback on retention of initially correct responses. Of course, we are not arguing against the fact that correcting memory errors is a key purpose of feedback. Instead, we believe that feedback also functions as an error-correction mechanism for correct responses, albeit for a different type of error. When individuals make a correct response but are not confident in the response, there is a discrepancy between the subjective and objective correctness of their answers. In other words, low-confidence correct responses reflect an error of metacognitive monitoring, which in this context refers to the ability to assess the accuracy of one's own performance on a test (Barnes, Nelson, Dunlosky, Mazzoni, & Narens, 1999; Koriat & Goldsmith, 1996; Nelson & Narens, 1990). Feedback that confirms the correctness of low-confidence responses should enable learners to reduce the discrepancy between their perceived and actual performance by allowing them to adjust their subjective assessments of their knowledge. Further, if feedback allows learners to correct initial metacognitive errors, then it should also enhance long-term retention of the correct responses and improve the accuracy of metacognitive monitoring on subsequent tests. Thus, our hypothesis in this research was that, just as feedback helps correct memory errors, feedback will also help correct meta-cognitive errors and will improve retention of low-confidence correct responses…*

Our hypothesis is that feedback serves to correct the metacognitive error inherent in low-confidence correct responses, much like it does for high-confidence errors in the hypercorrection effect. However, we believe that the correction of these two types of metacognitive error probably leads to better retention through different mechanisms. As described above, retention may be enhanced following high-confidence errors because a feeling of surprise causes subjects to pay more attention to feedback (Butterfield & Metcalfe, 2006). In contrast, we think that providing feedback after low-confidence correct responses might enhance retention by enabling learners to strengthen the association between the cue and response and to inhibit any competing responses…

the current experiments provide clear evidence that low-confidence correct responses do benefit from feedback and that feedback improves students' metacognitive judgments about their knowledge. Taken together, the two novel findings support the idea that a low-confidence correct response represents an error in metacognitive monitoring that can be corrected through feedback. Providing feedback after low-confidence correct responses enables learners to eliminate the discrepancy between perceived and actual correctness of the response. Feedback after both correct and incorrect responses on tests is a critical aspect of learning."

Butler, Karpicke & Roediger 2008

[192] *"Although metacognitive abilities are often treated as stable characteristics of individuals (Allen et al., 2017; Fleming, Weil, Nagy, Dolan, & Rees, 2010; McCurdy et al., 2013; Song et al., 2011), several lines of research hint at their malleability. For instance, practicing meditation boosts the accuracy of retrospective confidence judgments about recognition memory decisions (Baird, Mrazek, Phil- lips, & Schooler, 2014) and monitoring of decision errors can be modulated by drugs (Hester et al., 2012; Hauser et al., 2017) and brain stimulation (Harty et al., 2014). Moreover, recent work has identified distinct neural substrates in the frontal and parietal*

lobes supporting metacognitive monitoring across a range of tasks (Al- len et al., 2017; Baird, Smallwood, Gorgolewski, & Margulies, 2013; Cortese, Amano, Koizumi, Kawato, & Lau, 2016; Fleming et al., 2010; McCurdy et al., 2013; see Fleming & Dolan, 2012, for a review), suggesting the potential for targeted modulation of metacognition independently of changes in first-order performance. Previous attempts to improve metacognitive ability (confidence calibration) through explicit instruction, practice, feedback, or a combination of these manipulations have led to mixed results, with some studies documenting increases, and others documenting null findings (e.g., Adams & Adams, 1958; Bol, Hacker, O'Shea, & Allen, 2005; Lichtenstein & Fischhoff, 1980; Nietfeld & Schraw, 2002; Renner & Renner, 2001; Sharp, Cutler, & Penrod, 1988). One potential explanation for such heterogeneity of results is that training may impact first-order performance, thus masking subtle changes in metacognition because they are positively correlated (Fleming & Lau, 2014; Sharp et al., 1988)."

Carpenter et al. 2019

[193] *"…it remains poorly understood as to whether metacognition relies on a domain-general resource that is 'applied' to the task at hand, or whether different metacognitive processes are engaged when evaluating performance in different domains."*

Rouault et al. 2018

[194] *"That we as humans do make errors in thinking, judgment, and memory is undisputed. In fact, there is a plethora of phenomena showing that we deviate in our thinking, judgment, and memory from the objective and arguably 'correct' standard'."*

Pohl 2017

[195] *"Most researchers in social cognition favor the view that heuristics are simple, efficient shortcuts applied in judgment and decision-making when people face overly complex tasks, have limited time or cognitive ability, or deal with incomplete information in the world. In this light, heuristics work well in many instances, but are prone to break down in systematic ways-and whenever they do, more 'evidence' has been found that the mind is flawed in its reasoning abilities. The traditional treatment of heuristics has been largely dominated by researchers working within the heuristics and biases program (Kahneman, Slovic, & Tversky, 1982; Tversky & Kahneman, 1974), who have argued that human judgment often substantially deviates from optimality predictions or normative standards of logic."*

Haselton et al. 2009

[196] *"According to Gigerenzer, the collection of biases has focused too much on the few faulty cases of judgment and decision making thereby ignoring the majority of heuristics where heuristics typically lead to correct or at least useful decisions. This one-sided view may have led some researchers to conclude that 'mental illusions should be considered the rule rather than the exception' (Thaler 1991) and that 'mistakes of reason rule our mind' (Piattelli-Palmarini 1994). But this view of human rationality appears overly pessimistic…*

Gigerenzer (1991) has also asserted that non everything that looked like a cognitive illusion really is one. More specifically he argued that one could make cognitive illusions disappear by (a) avoiding too narrow norms to evaluate human performance, by (b) using an adequate statistical format, and by © using representative (instead of selected) samples of items. …

Using adequate statistical formats and representative sampling, Gigerenzer (1991; see also Gigerenzer & Hoffrage 1995) accordingly showed that some illusions could be substantially reduced or even eliminated, namely, the conjunction fallacy, … baserate neglect, … and overconfidence.

Pohl 2017

[197] *"A fundamental criticism of the heuristics and biases program is that researchers might be neglecting the structure of the world in which the decision-making takes place. Ecologically-minded scientists have argued that in order to understand the mind's true cognitive abilities one needs to consider the environment in which it operates-or was designed to operate by natural selection. For example, Egon Brunswik (1955) emphasized that psychologists should study how the mind makes inferences based on the informational cues present in the natural environment, and Roger Shepard (2001) saw the mind as a mirror reflecting regularities of the physical world (see Todd & Gigerenzer, 2007). Consequently, what we call a good or a bad decision (or rational and irrational behavior) has to be judged with regard to specific decision environments rather than in a vacuum (Gigerenzer, Todd,& the ABC Research Group, 1999)…*

A related criticism addresses how much information is usually available in these decision environments and if the clear standard for comparing decision outcomes-the supposedly optimal way of thinking-should necessarily be informed by abstract standards of probability, logic, and mathematical optimization. Many traditional models of rational choice assume that humans (and animals) make inferences about the world virtually as if they were supernatural beings that have unlimited reasoning power, boundless knowledge, and unlimited time to make their decisions. However, real-world decision environments, both current and past, do not look like this and it is unrealistic to compare the human capacity for judgment and choice against such optimality predictions and assumptions (see Gigerenzer et al., 1999). For example, humans almost never have access to all of the pertinent information needed for making a decision about which mate to choose, what foods are best to eat, or which house to buy. Rather than following models of unbounded rationality, researchers pointed out that many decisions are made in a boundedly rational way (i.e., under conditions of limited time, information, and cognitive processing) and that heuristics are psychologically plausible solutions in situations where the one best solution does not exist or cannot be reached anyhow."

Haselton et al. 2009

[198] *"From an evolutionary perspective, however, it would be surprising if the mind were really so woefully muddled. The mind is an intricate, evolved machine that has allowed humans to inhabit and exploit an incredible range of environments. Humans effectively solve a variety of social-ecological problems including large-scale cooperation, social exchange, habitat formation, agriculture, and cumulative culture. We are a remarkably intelligent species, capable of surviving and reproducing in a complicated and ever-changing world. Could it really be that the human mind is as deeply flawed as the literature suggests?"*

Haselton et al 2009

[199] *"The term 'cognitive illusion' has evolved in analogy to the better-known domain of 'optical illusions' (see Roediger 1996). The first and main feature of a phenomenon to count as an illusion thus is that it leads to a perception, judgment or memory that reliably deviates from 'reality'. IN cases of optical and memory illusions, it may be immediately evident what constitutes reality (because subjective perception and recall can be compared to external and original stimuli, respectively), but in thinking and judgment, the matter is less clear (Gigerenzer 1996). The problem concerns how to define an objectively 'correct' judgment or decision…*

As a second criterion, the observed phenomenon needs to deviate from the normative standard in a systematic fashion (i.e. in a predictable direction) rather than just randomly. Therefore, most designs include a control group, assuming that any deviations in the control group's data result from random error alone, while the experimental group shows in addition a systematic effect…

the mechanisms eventually leading to cognitive illusions typically include a number of probabilistic processes so that an illusion will not necessarily be observed on each and every single trial, but may only become evident as a systematic bias if the data are summed across a large number of trials or participants. A third aspect of illusions is that they appear involuntarily, that is, without specific instructions or deliberate will. They just happen. … This does not mean that motivational factors or conscious metacognitions may not be influential, too, but they are not the ultimate cause of the illusion itself. They only moderate its size…

Another aspect is that persons who have fallen prey to a cognitive illusion usually don't realize what has happened. … That is, illusioned persons are still convinced to have judged, decided, or recalled something to the best of their knowledge. As a consequence, and this constitutes the fourth cornerstone of the proposed definition, and illusion is hard if not impossible to avoid. While this is probably true for optical illusions, the criterion is much weaker for cognitive ones. For some illusions, proper instruction, careful selection of the material, or other procedural variations may reduce or even eliminate the illusion (as an example see Gigerenzer, Hertwig, Hoffrage & Sedlmeier 2008), while for other illusions, most (if not all) attempts to overcome the effect have failed (as an example, see Pohl & Hell 1996…). And finally, as the fifth point to consider and to distinguish cognitive illusions from other forms of typical errors, misunderstandings or faulty memories, illusions often appear as rather distinct from the normal course of information processing. An illusion somehow 'sticks out' as something special that 'piques our curiosity' (as Roediger 19996 put it) and thus attracts researchers to explain this unexpected but robust finding. In other words, ordinary forms of forgetting (leading to commission errors), or deviations resulting from simple misunderstandings would not be considered 'illusions'. … This is not to say that an illusion cannot be explained with ordinary and general mechanisms of information processing. In fact, one of the theoretical goals of research in cognitive illusions is to avoid the assumption of any special mechanisms that is responsible only for this one phenomenon, but instead to explain the observed effects with what one already knows about cognitive processes in general."

Pohl 2017

[200] *"On the surface, cognitive biases appear to be somewhat puzzling when viewed through an evolutionary lens. Because they depart from standards of logic and accuracy, they appear to be design flaws instead of examples of good engineering. Cognitive traits can be evaluated according to any number of performance criteria—logical sufficiency, accuracy, speed of processing, and so on. The value of a criterion depends on the question the scientist is asking. To the evolutionary psychologist, however, the evaluative task is not whether the cognitive feature is accurate or logical, but rather how well it solves a particular problem, and how solving this problem contributed to fitness ancestrally. Viewed in this way, if a cognitive bias positively impacted fitness, it is not a design flaw—it is a design feature."*

Haselton, Nettle & Murray 2015

[201] *"On the surface, cognitive biases appear to be somewhat puzzling when viewed through an evolutionary lens. Because they depart from standards of logic and accuracy, they appear to be design flaws instead of examples of good engineering. Cognitive traits can be evaluated according to any number of performance criteria—logical sufficiency, accuracy, speed of processing, and so on. The value of a criterion depends on the question the scientist is asking. To the evolutionary psychologist, however, the evaluative task is not whether the cognitive feature is accurate or logical, but rather how well it solves a particular problem, and how solving this problem contributed to fitness ancestrally. Viewed in this way, if a cognitive bias positively impacted fitness, it is not a design flaw—it is a design feature."*

Haselton, Nettle & Murray 2015

[202] "*By cognitive bias, we mean cases in which human cognition reliably produces representations that are systematically distorted compared to some aspect of objective reality." ...*

"*An evolutionary psychological perspective predicts that the mind is equipped with function-specific mechanisms adapted for special purposes—mechanisms with special design for solving problems such as mating, which are separate, at least in part, from those involved in solving problems of food choice, predator avoidance, and social exchange (e.g., Kenrick, Neuberg, Griskevicius, Becker, & Schaller, 2010). In the evaluation of cognitive biases, demonstrating domain specificity in solving a particular problem is a part of building a case that the trait has been shaped by selection to perform that function.*" ...

"Some design features that appear to be flaws when viewed in one way are revealed to be adaptations when viewed differently. If one were to only consider the idea that selection favors the maximization of direct reproductive success, for example, the fact that human females lose reproductive capability many years before death would appear a design flaw. However, there is evidence that women in traditional societies can enhance their inclusive fitness by transferring investment to their daughters' daughters as soon as the latter are of reproductive age (Voland & Beise, 2002). Viewed in this light, female menopause might be very well designed (Hawkes, 2003). In sum, there may be many evolutionary reasons for apparent design flaws, and a close examination often provides insight into the evolutionary forces that shaped them and their functions."

Haselton, Nettle & Murray 2015

[203] "*Knowledge of biases and illusions is of course valuable. For example, demonstrating that a bias may occur in some situations but not others (context effects), or with certain classes of information and not others (content effects), can reveal structural features of the mind. Additionally, knowledge of biases and illusions may have important practical utility by preventing undesirable outcomes.*"

Haselton et al 2009

[204] "*First, selection may favor useful shortcuts that tend to work in most circumstances, though they fall short of some normative standards (heuristics); second, apparent biases can arise if the task at hand is not one for which the mind is designed (artifacts); and third, biases can arise if biased response patterns to adaptive problems resulted in lower error costs than unbiased response patterns (error management biases).*"

Haselton, Nettle & Murray 2015

[205] "*Tversky and Kahneman attributed these and other biases to the operation of mental shortcuts : 'People rely on a limited number of heuristic principles which reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations' (1974, p. 1124). The gambler's fallacy and the conjunction fallacy are attributed to one of the most commonly invoked heuristics, representativeness, or the way in which A*

resembles or is representative of B. According to this account, alternating heads and tails are more representative of randomness than are series containing runs." …

"*Overall, there is ample evidence of cognitive bias and error in humans. Some of these biases may result from the use of shortcuts, which are often effective. For these effects, however, it is important to note that a 'processing limitations' explanation is not complete. Of all possible equally economical cognitive shortcuts, why were these particular ones favored by selection?"*

Haselton, Nettle & Murray 2015

[206] "*One type of artifact arises from evolutionarily novel problem formats. Gigerenzer (1997) proposed that tasks intended to assess human statistical prediction should present information in frequency (rather than probability) format, given that natural frequencies, such as the number of times an event has occurred in a given time period, are more readily observable in nature. In contrast, probabilities (in the sense of a number between 0 and 1) are mathematical abstractions beyond sensory input data, and information about the base rates of occurrence is lost when probabilities are computed (Cosmides & Tooby, 1996). Bayesian calculations involving frequencies are therefore computationally simpler than equivalent calculations involving probabilities, relative frequencies, or percentages. Whereas probability calculations need to reintroduce information about base rates, frequency calculations do not since this part of the computation is already 'done' within the frequency representation itself (Hoffrage, Lindsey, Hertwig, & Gigerenzer, 2001)* …

"A second artifact can arise from evolutionarily novel problem content. The perspective on cognitive design we have described suggests that researchers should not necessarily expect good performance in tasks involving abstract rules of logic. Falsification- based logic is sufficiently difficult for humans that university courses in logic, statistics, and research design attempt to teach it to students (with only mixed success)…

Haselton, Nettle & Murray 2015

[207] "*Within this framework, many ostensible faults in human judgment and evaluation may reflect the operation of mechanisms designed to make inexpensive, frequent errors rather than occasional disastrous ones (Haselton & Nettle, 2006; Johnson et al., 2013)* ...

"*Food Aversions Lasting aversion to a food is reliably acquired, in humans and other species, following a single incidence of sickness after ingestion of the food (Garcia, Hankins, & Rusiniak, 1976; Rozin & Kalat, 1971). Given one data point (sickness following the food type on one occasion), the system treats the food as if it is always illness-inducing. There are again two possible errors here. The false positive may be inconvenient, but the false negative is more likely to be fatal. The system appears biased toward over-responsiveness to avoid illness*...

"*The error management account is similar to that for food aversions: The false negative (failing to avoid someone with a contagious disease) is highly costly, whereas the false positive (avoiding contact with a non contagious person) may have small social or interpersonal costs, but is unlikely to have significant negative fitness consequences.*"

Haselton, Nettle & Murray 2015

[208] "*We ultimately conclude that the mind is best described as adaptively rational. By adaptably rational we mean that the mind shows evidence of psychological design for coping with recurrent adaptive problems our ancestors encountered over evolutionary history - the mind is equipped with mechanisms that are constrained and sometimes imprecise, but nevertheless clear products of natural selection showing evidence of good design. This definition runs in contrast to the often implicit definitions of rationality used by many social science researchers, including that the mind should maximize 'accuracy,' happiness, well-being, financial return, or adherence to abstract rules of logic. We do not deny that it is useful to compare human performance to these standards, as they may be those we wish to maximize in the modem world. Instead, we challenge the idea that deviations of performance from the standards means that the human mind is deeply flawed or poorly designed.*"

Haselton et al 2009

[209] "*Dual process models of higher cognition have become very influential in the cognitive sciences. The popular Default-Interventionist model has long favored a serial view on the interaction between intuitive and deliberative processing (or System 1 and System 2). Recent work has led to an alternative hybrid model view in which people's intuitive reasoning performance is assumed to be determined by the absolute and relative*

strength of competing intuitions. In the present study, we tested unique new predictions to validate the hybrid model…

There are different types of dual process models but arguably the dominant framework has been the serial or default-interventionist model that has been put forward by prominent scholars such as Daniel Kahneman (Kahneman, 2011) or Jonathan Evans and Keith Stanovich (Evans & Stanovich, 2013). At the core of the default-interventionist (DI) model lays a serial view on the interaction between System 1 and 2. The key idea is that when people are faced with a reasoning problem, they will typically rely on the fast System 1 to generate an answer. This is the default system. If needed, people can activate System 2 in a later phase to intervene and correct System 1 output (e.g., Evans, 2008; Evans & Stanovich, 2013). But this System 2 engagement only occurs after System 1 has been engaged and is not guaranteed. More generally, in the serial DI model, reasoners are conceived as cognitive misers who try to minimize cognitive effort (Kahneman, 2011). Since System 2 thinking is hard, people will often refrain from recruiting it and stick to the default System 1 response. The serial DI model offers an appealing explanation for the widespread 'bias' that has been observed in the reasoning and decision-making literature…

Why do intelligent adults so often violate basic logico-mathematical principles? Default-Interventionist theorists have highlighted that the key problem is that taking these principles into account typically requires demanding System 2 computations (e.g., Kahneman, 2011; Stanovich & West, 2000). When the fast System 1 has provided us with a response, most reasoners will refrain from engaging the effortful System 2. Consequently, they will not detect that their answer conflicts with more logical considerations (Kahneman, 2011). Put differently, these biased System 1 reasoners do not detect that they are being biased. The few people who do engage System 2 will override the initially cued intuitive System 1 response after their System 2 deliberation is completed and manage to give the correct logico-mathematical response."

Bago & De Neys 2019

[210] *"The default-interventionist model and the corresponding bias blind spot and corrective assumptions have had far-reaching impact on theorizing in the various fields that have adopted dual process models and, more generally, our view of human rationality (e.g., Gürçay & Baron, 2017; Stanovich & West, 2000). However, in recent years direct experimental testing of the core assumptions has pointed to fundamental issues. Pace the 'bias blind spot' hypothesis, a range of studies have established that often biased reasoners do show bias sensitivity (e.g., Bonner & Newell, 2010; De Neys & Glumicic, 2008; Gangemi, Bourgeois-Gironde, & Mancini, 2015; Pennycook, Trippas, Handley, & Thompson, 2014; Stupple, Ball, Evans, & Kamal-Smith, 2011; but see also Aczel, Szollosi, & Bago, 2016; Mata, Ferreira, Voss, & Kollei, 2017; Travers, Rolison, & Feeney, 2016). In these studies, participants are presented with both traditional reasoning problems in which a cued*

Bago & De Neys 2019

[211] *"The present study focuses on this critical error sensitivity issue. Preschool children were given a classic version of a number conservation task in which an intuitively cued response conflicted with the correct conservation response and a control version in which this conflict was not present. After solving each version children were asked to indicate their response confidence. Results showed that in contrast with children who gave a correct conservation response, preschoolers who erred showed a sharp confidence decrease after solving the classic conflict problem. This suggests that non-conserving preschoolers detect that their response is questionable and are less ignorant about conservation than their well-documented errors might have previously suggested."*

De Neys, Lubin & Houdé 2014

[212] *These findings point to at least two fundamental implications about the way we conceive intuitive and deliberate thinking or System 1 and System 2. On the one hand, it suggests that we might need to upgrade our view of the intuitive System 1. Although System 1 can frequently cue incorrect intuitions, it also generates correct intuitions. Among correct responders these correct intuitions will often dominat …*

The results of our justification studies suggest that instead of incorrection, the contribution of deliberate processing in these cases might rather lie in its cognitive transparency (Bonnefon, 2016). We observed that whereas people don't manage to explain why their initial response is correct, they seem to have little difficulties in providing such correct justifications after deliberation. Clearly, being able to produce a proper justification for one's insights is quite crucial…

Bago & De Neys 2019

[213] *"Executive functions (EFs) consist of a family of three, interrelated core skills (inhibitory control, working memory, and cognitive flexibility; Miyake et al., 2000, Diamond, 2013). From those, higher-order EFs are built such as reasoning, problem-solving, and planning (Collins and Koechlin, 2012, Lunt et al., 2012). Inhibitory control involves resisting one's initial impulse or a strong pull to do one thing, and instead act more wisely. Without inhibitory control we would be at the mercy of external stimuli, internal impulses, and habits of thought or action that pull us this way or that. Inhibitory control thus makes it possible for us to choose how we react and to change how we behave rather than being 'unthinking' creatures of habit or impulse (Diamond, 2013). It is critical for avoiding social faux pas and for a civil society where people abide by rules and norms. It is difficult to think of any aspect of life where having the presence of mind to wait before speaking or acting, giving a considered response rather than an impulsive one, being able to stay focused despite distraction, and resisting temptations to do inappropriate, ill-advised, self-destructive or illegal things would not be beneficial. Working memory (WM) involves more than holding information in mind. It involves doing that while performing one or more mental operations. It is needed, for example, for re-ordering the items you are holding in mind or seeing how they relate to one another ('working with' the information you are holding in mind; Baddeley and Hitch, 1994, Smith and Jonides, 1999) and also for remembering your question or comment while following an ongoing discussion or for holding in mind what you were about to do when something arises that must be dealt with first (D'Esposito and Postle, 2015). WM is critical for reasoning and problem-solving for they require holding lots of information in mind, exploring their interrelations, and then perhaps dis-assembling those combinations and re-combining the elements in new ways. WM is necessary for making sense of anything that unfolds over time for that always requires holding in mind what happened earlier and relating that to what is happening now (e.g., following a lecture or conversation, relating what you are reading now to what you read earlier, or understanding the relation between a later effect and an earlier cause).*

Cognitive flexibility refers to the ability to flexibly adjust to changed demands or priorities, to look at the same thing in different ways or from different perspectives (as required for set shifting or task switching; Allport and Wylie, 2000, Kiesel et al., 2010, Monsell, 2003, Vandierendonck et al., 2010). If one way of solving a problem isn't working, one needs cognitive flexibility to 'think outside the box,' that is, to find other ways of conceiving of the problem or of attacking it. Such flexibility is needed for meeting novel, unanticipated challenges and for seizing opportunities when they unexpectedly arise.

EFs are predictive of achievement, health, wealth, and quality of life throughout life, often more so than IQ or socioeconomic status (SES; e.g., Moffitt et al., 2011, Moffitt, 2012). They are more critical for school readiness than IQ or entry-level reading or math (Alloway et al., 2005, Blair, 2002, Blair and Razza, 2007, Carlson and Moses, 2001, Hughes and Ensor, 2008, Morrison et al., 2010). They are predictive of success throughout the school years from preschool through university (often more so than IQ [Duckworth and Seligman, 2005, Alloway and Alloway, 2010, Borella et al., 2010, Duncan et al., 2007, Fiebach et al., 2007, Gathercole et al., 2004, Loosli et al., 2012, McClelland et al., 2007, Nicholson, 2007, Savage et al., 2006])."

Diamond & Ling 2016

[214] "*Recent studies reported that training of working memory may improve performance in the trained function and beyond. Other executive functions, however, have been rarely or not yet systematically examined. The aim of this study was to test the effectiveness of inhibitory control (IC) training to produce true training-related function improvements in a sample of 122 healthy adults using a randomized, double-blind pretest/posttest/follow-up design. Two groups performed either adaptive (training group) or non-adaptive (active control) versions of go/no-go and stop-signal tasks for 3 weeks. Training gains as well as near-transfer to an untrained Stroop task and far-transfer to psychometric fluid intelligence were explored. Although the adaptive group could substantially improve overall IC task performance after training, no differences to the active control group occurred, neither at posttest nor at follow-up testing. A large decrease in response latency from pre- to posttest (and from pretest to 4 months' follow-up testing) was found when the training group was compared to the passive control group, which, however, does not sufficiently control for possible confounds. Thus, no conclusive evidence was found that this performance increase mirrors a true increase in IC function. The fact that training improvement was mainly related to response latency may indicate that individuals were more focused on performance gains in the prepotent go trials but less on the stop trials to meet the requirements of the tasks as well as possible. The challenges for response inhibition training studies are extensively discussed. (PsycINFO Database Record (c) 2019 APA, all rights reserved).*"

Enge et al. 2014

[215] "*Objectives: Inhibitory control training has been hypothesised as a technique that will improve an individual's ability to overrule impulsive reactions in order to regulate behaviour consistent with long-term goals. Methods: A meta-analysis of 19 studies of inhibitory control training and health behaviours was conducted to determine the effect of inhibitory control training on reducing harmful behaviours. Theoretically driven moderation analyses were also conducted to determine whether extraneous variables account for heterogeneity in the effect; in order to facilitate the development of effective intervention strategies. Moderators included type of training task, behaviour targeted, measurement of behaviour and training duration. Results: A small but homogeneous effect of training on behaviour was found,* $d^+ = 0.378$, $CI_{95} = [0.258, 0.498]$. *Moderation analyses revealed that the training paradigm adopted, and measurement type influenced the size of the effect such that larger effects were found for studies that employed go/no-go (GNG) training paradigms rather than stop-signal task paradigms, and objective outcome measures that were administered immediately yielded the largest and most consistent effects on behaviour. Conclusions: Results suggest that GNG inhibitory control training paradigms can influence health behaviour, but perhaps only in the short-term. Future research is required to systematically examine the influence of training duration, and the longevity of the training effect. Determining these factors could provide the basis for cost-effective and efficacious health-promoting interventions.*"

Allom, Mullan & Hagger 2016

[216] "*The 'Executive Functions' (EFs) of inhibitory control, working memory, and cognitive flexibility enable us to think before we act, resist temptations or impulsive reactions, stay focused, reason, problem-solve, flexibly adjust to changed demands or priorities, and see things from new and different perspectives. These skills are critical for success in all life's aspects and are sometimes more predictive than even IQ or socioeconomic status. Understandably, there is great interest in improving EFs. It's now clear they can be improved at any age through training and practice, much as physical exercise hones physical fitness. However, despite claims to the contrary, wide transfer does not seem to occur and 'mindless' aerobic exercise does little to improve EFs. Important questions remain: How much can EFs be improved (are benefits only superficial) and how long can benefits be sustained? What are the best methods for improving EFs? What about an approach accounts for its success? Do the answers to these differ by individual characteristics such as age or gender? Since stress, sadness, loneliness, or poor health impair EFs, and the reverse enhances EFs, we predict that besides directly train EFs, the most successful approaches for improving EFs will also address emotional, social, and physical needs.*"

Diamond & Ling 2016

[217] "*To provide evidence for the role of inhibitory control in overcoming deductive reasoning errors, we contrasted the effect of two types of training on the ability to perform deductive reasoning tasks. In one condition, participants were trained to inhibit the perceptual matching bias. In the other condition, participants received training focusing on explaining the underlying logic of the task. Importantly, participants were trained on a different deductive task (i.e., the Wason task, Wason, 1968) than the one performed pre- and post-training (i.e., the perceptual matching bias task, Evans, 1998). The effects of the two types of training were compared to a test-retest control condition in which participants simply performed the perceptual matching task two times. Participants who were trained to inhibit the perceptual matching heuristic were the only ones who succeeded to overcome their deductive reasoning errors. This finding suggests that logical reasoning errors are not due to a lack of logic (or experience) but to a default to inhibit a misleading heuristic. In a follow-up PET (positron emission tomography) imaging study in which we compared the cerebral activation before and after the participants were trained in inhibiting the perceptual matching bias, we observed that the brain activation shifted from the posterior perceptual regions pre-training to prefrontal executive regions post-training. This is the first micro-longitudinal neuroimaging study of deductive reasoning and it provides the first evidence that inhibitory control was critical to reason logically. Note that this brain imaging study on reasoning errors correction was conducted on a sample of only eight participants but the strength of these results stem from the fact that the participants were their own controls in the pre-post training comparison.*"

Houdé & Borst 2015

[218] "*Training interventions to improve decision making, to date, have met with limited success mostly in specific domains. Training can be very effective when accuracy requires experts to recognize patterns and select an appropriate response, such as in weather forecasting, firefighting, and chess (Phillips, Klein, & Sieck, 2004). By contrast, even highly trained professionals are less accurate than very simple mathematical models in other domains such as parole decisions, personnel evaluations, and clinical psychological testing (Dawes, Faust, &*

Meehl, 1989). Whether domain- specific expertise is achievable appears to be contingent on external factors such as the prevalence of clear feedback, the frequency of the outcome being judged, and the number and nature of variables that determine that outcome (Harvey, 2011; Kohler, Brenner, & Griffin, 2002).

Evidence that training effectively improves general decision-making ability is inconclusive at present (Arkes, 1991; Milkman, Chugh, & Bazerman, 2009; Phillips et al., 2004). Weather forecasters are well calibrated when predicting the chance of precipitation (Murphy & Winkler, 1974), for example, but are overconfident in their answers to general knowledge questions (Wagenaar & Keren, 1986). Even within their domain of expertise, experts struggle to apply their training to new problems. Philosophers trained in logic exhibit the same preference reversals in similar moral dilemmas as academics without logic training (Schwitzgebel & Cushman, 2012), and physicians exhibit the same preference reversals as untrained patients for equivalent medical treatments when those treatments are framed in terms of survival or mortality rates (McNeil, Pauker, Sox, & Tversky, 1982). Several studies have shown that people do not apply their training to unfamiliar and dissimilar domains because they lack the necessary metacognitive strategies to recognize underlying problem structure (for reviews, see Barnett & Ceci, 2002; Reeves & Weisberg, 1994; Willingham, 2008)."

Merewedge et al. 2015

[219] "*From what we have learned about decision debiasing it seems that it is not enough to educate people about the existence of biases and their functionality; they also need to acquire specific debiasing strategies to cope with these challenges. What makes debiasing even more difficult is that decision makers have to recognize the situations in which they need to use the strategies they learned. This task requires transfer from the rule they learned during the training event to the test situation or (preferably) to any analogous real-life situation. The question is how to train people on an abstract rule that they would apply in various relevant situations. In studies of reasoning, some evidence indicates that practicing only abstract rules can improve performance on specific problems. For example, Fong et al. (1986) found that after training on the law of large numbers where the participants were taught about statistical notions such as sample, population and variability, they were better at reasoning about various uncertainty-related problems, such as slot machines, lotteries, or athletic performance. Similarly, undergraduate and graduate training in psychology and social sciences (Lehman et al., 1988; Lehman and Nisbett, 1990) has been found to increase the students' ability in reasoning about everyday problems involving uncertainty. Nevertheless, it is hard to assume that these people relied only on their abstract knowledge for the new cases and did not benefit from the concrete examples (e.g., the urn problem demonstrations in the study of Fong et al., 1986) use during the training. Closer examination of these results suggests that it is easier to apply the abstract rules in cases with matching superficial features. For example, Cheng et al. (1986) showed that abstract training of the obligation rule ('If precondition P is satisfied, action A must be taken') improves performance on Wason's (1966) four-card problem, but only on those versions of the task where the obligation rule could be used in the task. Fong and Nisbett (1991) taught their participants about the law of large numbers in either one of two domains and they were tested on both domains. Although immediately after the training, they found no effect of domain, 2 weeks later the participants could perform better in the domain they were taught in. In their summary, Smith et al. (1992) suggest that in situations where more than one mechanism is involved, reasoning might rely on hybrid instance-rule mechanisms. Therefore, superficial similarity between the learned instance and the target case can facilitate rule-application. This suggestion is in accord with studies of problem solving where it is assumed that a major cause of failures to transfer the relevant rule to analogous situations is the greater attention people pay to the salient and superficial details at the time of learning and that they will apply the learned principles in the test situation to the degree that it shares those contextual features (Holyoak and Koh, 1987; Ross, 1987).*"

Aczel et al. 2015

[220] "*In each case, expertise in a domain helps people develop a sensitivity to patterns of meaningful information that are not available to novices… For example, electronics technicians were able to reproduce large portions of complex circuit diagrams after only a few seconds of viewing ; novices could not. The expert circuit technicians chunked several individual circuit elements (eg. resistors and capacitors) that performed the function of an amplifier. By remembering the structure and function of a typical amplifier, experts were able to recall the arrangement of the many individual circuit elements comprising the '« amplifier chunk' ». … physicists recognize problems of river currents and problems of headwinds and tailwinds in airplanes as involving similar mathematical principles, such as relative velocities*." »

Bransford, Brown & Cocking 2000

[221] "*Experts notice features and meaningful patterns of information that are not noticed by novices.*

Experts have acquired a great deal of content knowledge that is organized in ways that reflect a deep understanding of their subject matter.
Expert knowledge cannot be reduced to sets of isolated facts or propositions but, instead, reflects contexts of applicability ; that is, the knowledge is '« conditionalized' » on a set of circumstances.
Experts are able to flexibly retrieve important aspects of their knowledge with little attentional effort.
Though experts know their disciplines thoroughly, this does not guarantee that they are able to teach others.
Experts have varying levels of flexibility in their approach to new situations."

Bransford et al. 2000

[222] « *…lorsque les connaissances conceptuelles des sujets sont contrôlées (ce qu'on nomme parfois le niveau d'expertise) relativement à des domaines précis (le football, les échecs, les dinosaures), il devient impossible de mettre en évidence un développement : ce qui en savent plus font mieux que les autres, quels que soient les âges (Chi et Ceci 1987 ; Yates & Chandler 1991).* »

Fayol & Monteil 1994

[223] « *(Les stratégies) sont censées fournir au sujet un ensemble d'outils adaptables, transférables, susceptibles de lui permettre d'apprendre à apprendre tout en résolvant les problèmes soulevés par la vie, scolaire d'abord, quotidienne ensuite*. »

Fayol & Monteil 1994

[224] « *Une stratégie est une séquence intégrée, plus ou moins longue et complexe, de procédures sélectionnées en vue d'un but afin de rendre optimale la performance. Elle peut concerner des procédures très générales – par exemple l'idée qu'il est nécessaire de planifier intentionnellement pour atteindre un but – ou très spécifiques – par exemple, se poser à soi-même des questions pour s'assurer qu'on a bien compris un texte (Nisbet & Shucksmith, 1986).* »

Fayol & Monteil 1994

[225] « *Le facteur le plus influent en ce qui concerne la mobilisation des procédures semble être la connaissance préalable du sujet relativement au domaine conceptuel évoqué*. »

Fayol & Monteil 1994

[226] "*Studies of the Philosophy for Children program may be taken as typical. Two researchers identified eight studies that evaluated academic outcomes and met minimal research-design criteria. (Of these eight, only one had been subjected to peer review.) Still, they concluded that three of the eight had identifiable problems that clouded the researchers' conclusions. Among the remaining five studies, three measured reading ability, and one of these reported a significant gain. Three studies measured reasoning ability, and two reported significant gains. And, two studies took more impressionistic measures of student's participation in class (e.g., generating ideas, providing reasons), and both reported a positive effect.*"

Willingham 2007

[227] "*Despite the difficulties and general lack of rigor in evaluation, most researchers reviewing the literature conclude that some critical thinking programs do have some positive effect. But these reviewers offer two important caveats. First, as with almost any educational endeavor, the success of the program depends on the skill of the teacher. Second, thinking programs look good when the outcome measure is quite similar to the material in the program. As one tests for transfer to more and more dissimilar material, the apparent effectiveness of the program rapidly drop*s."

Willingham 2007

[228] "*If you remind a student to 'look at an issue from multiple perspectives' often enough, he will learn that he ought to do so, but if he doesn't know much about an issue, he can't think about it from multiple perspectives. You can teach students maxims about how they ought to think, but without background knowledge and practice, they probably will not be able to implement the advice they memorize*."

Willingham 2007

[229] "*Can critical thinking actually be taught? Decades of cognitive research point to a disappointing answer: not really. People who have sought to teach critical thinking have assumed that it is a skill, like riding a bicycle, and that, like other skills, once you learn it you can apply it in any situation. Research from cognitive science shows*

that thinking is not that sort of skill. The processes of thinking are intertwined with the content of thought (that is, domain knowledge).”

Willingham 2007

[230] *“If knowledge of how to solve a problem never transferred to problems with new surface structures, schooling would be inefficient or even futile—but of course, such transfer does occur. When and why is complex, but two factors are especially relevant for educators: familiarity with a problem's deep structure and the knowledge that one should look for a deep structure.*”

Willingham 2007

[231] *“Here's an example: A treasure hunter is going to explore a cave up on a hill near a beach. He suspected there might be many paths inside the cave so he was afraid he might get lost. Obviously, he did not have a map of the cave; all he had with him were some common items such as a flashlight and a bag. What could he do to make sure he did not get lost trying to get back out of the cave later? The solution is to carry some sand with you in the bag, and leave a trail as you go, so you can trace your path back when you're ready to leave the cave. About 75 percent of American college students thought of this solution—but only 25 percent of Chinese students solved it.6 The experimenters suggested that Americans solved it because most grew up hearing the story of Hansel and Gretel, which includes the idea of leaving a trail as you travel to an unknown place in order to find your way back. The experimenters also gave subjects another puzzle based on a common Chinese folk tale, and the percentage of solvers from each culture reversed. (To read the puzzle based on the Chinese folk tale, and the tale itself, go to www.aft.org/pubs-reports/american_educator/index.htm.) It takes a good deal of practice with a problem type before students know it well enough to immediately recognize its deep structure, irrespective of the surface structure, as Americans did for the Hansel and Gretel problem.*”

Willingham, 2007

[232] *“They are little chunks of knowledge—like 'look for a problem's deep structure' or 'consider both sides of an issue'—that students can learn and then use to steer their thoughts in more productive directions.” Le problème de ce genre de stratégie métacognitive est qu'elle ne peut pas porter beaucoup plus loin qu'à amener à se répéter au bon moment la stratégie métacognitive. 'Thus, a student who has been encouraged many times to see both sides of an issue, for example, is probably more likely to spontaneously think “I should look at both sides of this issue” when working on a problem. … Unfortunately, metacognitive strategies can only take you so far. Although they suggest what you ought to do, they don't provide the knowledge necessary to implement the strategy.”*

Willingham 2007

[233] “*Understanding and using conditional probabilities is essential to scientific thinking because it is so important in reasoning about what causes what. But people's success in thinking this way depends on the particulars of how the question is presented. Studies show that adults sometimes use conditional probabilities successfully, but fail to do so with many problems that call for it. Even trained scientists are open to pitfalls in reasoning about conditional probabilities (as well as other types of reasoning). Physicians are known to discount or misinterpret new patient data that conflict with a diagnosis they have in mind, and Ph.D.- level scientists are prey to faulty reasoning when faced with a problem embedded in an unfamiliar context.*”

Willingham 2007

[234] “*But critical thinking is very different. As we saw in the discussion of conditional probabilities, people can engage in some types of critical thinking without training, but even with extensive training, they will sometimes fail to think critically. This understanding that critical thinking is not a skill is vital.*”

Willingham 2007

[235] “*It tells us that teaching students to think critically probably lies in small part in showing them new ways of thinking, and in large part in enabling them to deploy the right type of thinking at the right time.*”

“*What do all these studies boil down to? First, critical thinking (as well as scientific thinking and other domain-based thinking) is not a skill. There is not a set of critical thinking skills that can be acquired and deployed regardless of context. Second, there are metacognitive strategies that, once learned, make critical thinking more likely. Third, the ability to think critically (to actually do what the metacognitive strategies call for) depends on*

domain knowledge and practice. For teachers, the situation is not hopeless, but no one should underestimate the difficulty of teaching students to think critically."

Willingham 2007

[236] "*Numerous qualitatively different forms of outcome evaluations for thinking courses provide substantial evidence for the conclusion that it is possible to use education to improve the ability to think critically, especially when instruction is specifically designed to encourage transfer of these skills to different situations and different domains of knowledge*."

Halpern 2013

[237] "*Critical thinking does not automatically result as a byproduct of standard instruction in a content area. Critical thinking instruction needs to focus overtly and self-consciously on the improvement of thinking, and the learning experience needs to include multiple examples across domains in order to maximize transfer*."

Halpern 2013

[238] "*Ideally critical thinking skills should be used to recognize and resist unrealistic campaign promises, circular reasoning, faulty probability estimates, weak arguments by analogy, or language designed to mislead whenever and wherever it is encountered. Critical thinkers should be able to solve or offer reasonable solutions to real world problems, whether it is the problem of nuclear war or how to set up a new computer. These skills should also be long lasting and used for the many decades of critical thinking that most of us will face…*"

Halpern 2013

[239] "*The best way to promote the kind of transfer I am advocating is with the conscious and deliberate use of the skills that are learned in a wide variety of contexts.*"

Halpern 2013

[240] "*Our working definition for the purposes of this review is that thinking skills interventions are approaches or programmes which identify for learners translatable, mental processes and/or which require learners to plan, describe and evaluate their thinking and learning. These can therefore be characterised as approaches or programmes which:*

- *require learners to articulate and evaluate specific learning approaches; and/or*
- *identify specific cognitive, and related affective or conative processes that are amenable to instruction.*"

Higgins et al. 2005

[241] "*Studies were selected for the meta-analysis if they had sufficient quantitative data to calculate an effect size (relative to a control or comparison group of pupils) and if the number of research subjects was greater than 10. Effect sizes were calculated from the reported data and combined statistically using quantitative synthesis*." Le résultat est que "*Twenty-nine studies were identified which contained quantitative data on pupils' attainment and attitudes suitable for meta-analysis. The studies come from a range of countries around the world with half set in the US and UK. The studies broadly cover the ages of compulsory schooling (5–16) and include studies set in both primary and secondary schools. A number of named thinking skills interventions are included, such as Feuerstein's instrumental enrichment (FIE) and cognitive acceleration through science education (CASE) as well as studies which report a more general thinking skills approach (such as the development of metacognitive strategies).*"

Higgins et al. 2005

[242] "*Twentieth-century psychologists have been pessimistic context of solving recurrent everyday problems. These rule about teaching reasoning, prevailing opinion suggesting that people may possess only domain-specific rules, rather than abstract rules; this would mean that training a rule in one domain would not produce generalization to other domains. … We propose an alternative view that is close to the pre-20th- century one: people do make use of inferential rules and these can be readily taught. In fact, rules that are extensions of naturally induced ones can be taught by quite abstract means. This description does not apply to formal, deductive logical rules or to most other purely syntactic rule systems, however. Instead, the types of inferential rules that people use naturally and can be taught most easily are a family of pragmatic inferential rule systems that people induce in the context of solving recurrent everyday problems (4). These rule systems are abstract in as much as they can be used in a wide variety of content domains, but their use is confined to certain types of*

problem goals and particular types of relations between events. They include 'causal schemas' (5), 'contractual schemas,' such as the rules underlying permission and obligation in the social sphere, and 'statistical heuristics,' used in the evaluation of evidence, such as qualitative, intuitive versions of the law of large numbers."

Nisbett et al. 1986

[243] "*The data base students acquire in school ought to inform their thinking in other subjects and in life outside the school*."

Perkins & Salomon 1988

[244] "*Analogy and transfer of learning … are both concerned with the same basic questions: how can what has been learnt in a certain situation be generalized to new situations and what conditions are necessary and/or sufficient to generate transfer from one situation to anothe*r?"

Gamo, Sander & Richard 2010

[245] "*Since transfer between tasks is a function of the similarity by transfer trasks and learning experiences, an important strategy for enhancing transfer from schools to other settings may be to better understand the non-school environments in which students must function."*

Bransford, Brown & Cocking 2000

[246] "*The transfer literature suggests that the most effective transfer may come from a balance of specific examples and general principles, not from either one alone*."

Bransford et al. 2000

[247] "*Instead of worrying about which is more important – local knowledge or the more general, transferable aspects of knowledge – we should recognize the synergy of local and more general knowledge. … students who lack an understanding of key mathematical concepts will not gain much from the general strategy of trying to define and represent a problem well before they start. But students who lack the habit of trying to define and represent a problem well will often misuse the mathematical concepts they know when the problem is not routine. … Proper attention to transfer will make the best of both for the sake of deeper and broader skill, knowledge and understanding*."

Salomon & Perkins 1987

[248] "*Humans are not naturally critical. Indeed, like ballet, critical thinking is a highly contrieved activity. Running is natural; night club dancing is less so; but ballet is something people can only do well with many years of painful, expensive, dedicated training."*

Van Gelder 1995

[249] "*Thus critical thinking cannot be treated just as a kind of gloss on educational content made up to other 'real' subjects. Students will not become excellent critical thinkers merely by studying history, marketing or nursing, even if the instruction is given a 'critical' emphasis (as it should be). Critical thinking must be studies and practiced in its own right; it must be an explicit part of the curriculum.*"

Van Gelder 1995

[250] "*Critical thinking skills are, by definition, ones that apply in a wide range of domains, contexts, and so on, and there is plenty of territories in which they can fail to transfer*."

Van Gelder 1995

[251] "*There is little evidence of the benefit of teaching metacognitive approaches in 'learning to learn' or 'thinking skills' sessions. Pupils find it hard to transfer these generic tips to specific tasks. Self-regulated learning and metacognition have often been found to be context-dependent, so how you best plan in Key Stage 2 art may have significant differences to planning strategies in GCSE maths. This means that pupil who shows strong self-regulated learning and metacognitive competence in one task or subject domain may be weak in another, and metacognitive strategies may or may not be effective, depending on the specific task, subject, or problem tackled. This does not, however, mean that metacognitive knowledge and skills will automatically develop through content knowledge teaching. That being said, over time, metacognition can become more generic, and older*

metacognitive learners can possess an array of strategies that they then judiciously apply across a range of contexts and to a range of tasks. This maturation also includes the development of a growing understanding of when to use what strategies, or when good strategies may be missing in the learner's repertoire."

EEF, https ://educationendowmentfoundation.org.uk/tools/guidance-reports/metacognition-and-self-regulated-learning

[252] "*In contrast to approaches that emphasize explicit instruction as the most powerful tool in developing critical thinking and writing (Larson, 100 Britt, & Kurby, 2009; Marin & Halpern, 2011; Schworm & Renkl, 2007), our approach is experiential in its pedagogical emphasis and microgenetic with regard to research methodology (Kuhn, 1995). By observing students engaging in guided practice over a period of time, we believe we can learn something about what develops and how*."

Kuhn, Hemberger & Khait 2015

[253] "*Our earlier work (Felton & Kuhn, 2001; Kuhn & Udell, 2003; Kuhn, Goh, Iordanou, & Shaenfield, 2008) documented that young adolescents engaged in argumentation typically concentrate their attention on exposition of their own claims, essentially ignoring the opponent's position. Thus, the initial goals of our programme are to encourage attention to the other's position and to enhance ability and disposition to address it with the goal of weakening it, or in other words, to engage in counter argumentation. Once this goal is achieved, our emphasis shifts to the use of evidence to strengthen and weaken claims. By securing answers to their own self-generated questions, students begin to contribute to the set of evidence bearing on the topic that we provide initially, and this evidence comes to play an increasing role in their argumentation.*"

Kuhn, Hemberger & Khait 2015

[254] "*In their discourse with one another, students will not be able to generate rich arguments and counterarguments in a vacuum. They need to bring to bear information relevant to the topic to inform their reasoning. Depending on the topic, they will likely already have some degree of related knowledge that they are ready to summon to support their arguments (or to weaken those of their opponents). But they will need more knowledge than they have at the outset. One approach we might take is to ask students to begin their work on a topic by reading material about it that they can then draw on in their argumentation. A problem with this approach is that students don't yet appreciate the purpose that this information serves. In a word, it provides answers to questions they don't yet have. As a result, they fail to see its point. They are thus likely to approach such reading disinterestedly, as just another reading assignment to be completed and most likely forgotten. A small dose of initial reading can be productive, to heighten initial interest in the topic, but at the outset we employ it sparingly out of concern that a deluge of information up front not only is met with disinterest but can shut down students' own thinking and inquisitiveness about a topic. Therefore, we let students' own ideas dominate at the beginning of their engagement with a new topic, encouraging them to articulate and share with one another their ideas about the topic. And, we have found, they do have lots of ideas to share, even in the case of topics outside the range of their immediate experience that we might expect them to know little about. Still, students' discourse stands to be informed and enriched by information bearing on the topic that likely will be new to them. An effective way to do this, we have found, is to create a need for the information they acquire. Rather than provide answers to questions students don't have, we let them first formulate the questions. In this way, we allow students to first see how such information could be useful in achieving their discourse objectives, and then we assist them in securing it. Hence, after introducing a few basic questions and answers regarding the topic, we invite students to generate questions of their own, the answers to which they think might be helpful to them. By the next session, we then make available brief factual answers to these questions (which sometimes students assist in obtaining), and the resulting question-and-answer 'evidence set' (containing all questions and answers) we assemble for them as a written document that is made available for use by the entire class throughout their work on the topic. The point is for students not just to acquire information but to see its value and therefore be disposed to apply it. With practice, we have found, students do in fact in time make much use of this information, coming to recognize it as playing a critical role in their discourse.*"

Kuhn, Hemberger & Khait 2015

[255] "*According to the argumentative theory of reasoning, the main function of human reasoning is to exchange arguments with others (Mercier & Sperber, 2011). This theory predicts that reasoning should yield better outcomes when it is used in conversation than in solitary ratiocination. It is thus well supported by the*

comparison of individual and group performance on intellective tasks (as well as on a variety of others tasks, see Mercier, 2016; Mercier & Sperber, 2011)."

Boku, Yama & Mercier 2018

[256] "*However, for these results to fully support the argumentative theory of reasoning, two conditions must hold. First, it is the exchange of arguments that must be chiefly responsible for the improvement in performance during discussion and not, for instance, the forceful imposition of a view by a more confident participant. Second, the improvement in performance during discussion should not depend on a particular cultural setting. Regarding the first condition, substantial evidence suggests that it is the exchange of arguments that is responsible for the improvement in performance during discussion. Transcripts reveal that participants do exchange arguments at length (Moshman & Geil, 1998; Trognon, 1993). Moreover, a participant with the correct answer convinces her peers even if she faces group members who all agree on a different answer (the same within these group members) and are more confident than she is (Trouche, Sander, & Mercier, 2014).*"

Boku, Yama & Mercier 2018

[257] "*Finally, participants who change their mind following group discussion tend to perform better on transfer problems than they had before the discussion (Laughlin, 2011; Trouche et al., 2014). Although the argumentative theory of reasoning does not predict that such transfer will always take place, when it does it is a sign that participants have understood why they should change their mind to accept the correct answer.*"

Boku, Yama, Mercier 2018

[258] "*Ideally, the performance in the post-test would have been compared to that of a control condition in which participants would have solved the initial problems twice as well (at the pre-test and test), but always on their own. However, such tests have never revealed significant improvements in performance (see, Laughlin, 2011). Given the huge improvement in performance observed during the test (from 0.07% correct answer to 0.61%), it is extremely implausible that mere repetition could explain the effect. Moreover, repetition could not explain the application of the truth- wins scheme. As a result, we can be confident that it is the group phase that yielded the improved performance.*"

Boku, Yama, Mercier 2018

[259] "*The most straightforward consequence of the combination of myside bias and laziness that characterizes argument production is that when people reason on their own, reasoning typically fails to correct misguided intuitions (e.g., [24]). For instance, most participants who tackle the bat and ball on their own fail to provide the right answer, although it is mathematically trivial [25]. Solitary reasoning can even lead to an accumulation of arguments supporting our opinions, most of them poorly examined, leading to overconfidence [22, 23] and polarization [26]. Solitary reasoning also provides people with excuses to engage in morally dubious behavior [27].*"

Mercier 2016

[260] "*The argumentative theory of reasoning predicts that people evaluate others' arguments in a way that is objective and demanding. However, in contrast with argument production, which has immediately perceivable effects (the arguments produced), argument evaluation is more difficult to measure. The evaluation of the strength of an argument is only one of the factors that determine whether its conclusion is accepted. Prior beliefs regarding the plausibility of the argument's conclusion and the trustworthiness of the argument's source also have to be taken into account. As a result, an argument's conclusion might be rejected, not because the argument was not evaluated properly, but because it failed to generate enough conflict with one's prior beliefs to lead to a change of mind.*"

Mercier 2016

[261] "*The study of persuasion and attitude change has shown that when participants are given arguments on issues that they care about, good arguments are much more effective at changing their minds than weak ones [12]. When participants evaluate everyday arguments, they react appropriately to variations in argument strength, whether strength is measured by Bayesian modeling [49–52], norms of classical logic [53], or norms of argumentation fallacies [54–57]. For instance, arguments from authority are deemed potentially fallacious if they come from non-expert sources or from sources with conflicts of interest, and participants appropriately discount such arguments [55].*"

Mercier 2016

Références

1. Introduction

- Acerbi, A. (2019). Cognitive attraction and online misinformation. *Palgrave Communications*, 5(1), 15.
- Evans, J. S. B. T. (2012). Dual process theories of deductive reasoning: facts and fallacies. *The Oxford handbook of thinking and reasoning*, 115-133.
- Gelder, T. V. (2005). Teaching critical thinking: Some lessons from cognitive science. *College teaching*, 53(1), 41-4
- Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge university press.
- Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Disposition, skills, structure training, and metacognitive monitoring. *American psychologist*, *53*(4), 449.
- Halpern, D. F. (2013). *Thought and knowledge: An introduction to critical thinking*. Psychology Press.
- Howells, K. (2018). *The future of education and skills: education 2030: the future we want*.
- Kahneman, D. (2002). Maps of bounded rationality: A perspective on intuitive judgment and choice. *Nobel prize lecture*, *8*, 351-401.
- Kuhn, D. (1999). A developmental model of critical thinking. *Educational researcher*, *28*(2), 16-46.
- Nisbett, R. E. (2015). *Mindware: Tools for smart thinking*. Farrar, Straus and Giroux.
- Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin*, 28(3), 369-381.
- Shafir, E., & LeBoeuf, R. A. (2002). Rationality. *Annual review of psychology*, *53*(1), 491-517.
- Stanovich, K. E. (1999). *Who is rational?: Studies of individual differences in reasoning*. Psychology Press.
- Stanovich, K. E. (2009). *What intelligence tests miss: The psychology of rational thought*. Yale University Press.
- Stanovich, K. E., & Stanovich, P. J. (2010). *A framework for critical thinking, rational thinking, and intelligence. Innovations in educational psychology: Perspectives on learning, teaching and human development*, 195-237.
- Van Gelder, T. (2005). Teaching critical thinking: Some lessons from cognitive science. *College teaching*, 53(1), 41-48.
- Whitworth, A. (2009). *Information obesity*. Elsevier.

2. Analyse des approches

Références citées dans le texte

- Bailin, S. (2002). Critical thinking and science education. *Science & Education*, 11(4), 361–375.
- Bailin, S., & Battersby, M. (2016). *Reason in the balance: An inquiry approach to critical thinking*. Hackett Publishing.
- Bailin, S., Case, R., Coombs, J. R., & Daniels, L. B. (1999). Conceptualizing critical thinking. *Journal of curriculum studies*, *31*(3), 285-302.
- Beran, M. J., Perner, J., & Proust, J. (Eds.). (2012). *Foundations of metacognition*. Oxford University Press.
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain.* New York: David McKay Co Inc.
- Boisvert, J. (2016). La formation de la pensée critique: deuxième partie. Bulletin de La Documentation Collégiale N° 15, Janvier 2016.
- Bronner, G. (2013). *La démocratie des crédules*. Presses universitaires de France.
- Cosmides, L., & Tooby, J. (1994). Better than rational : Evolutionary psychology and the invisible hand. *The American Economic Review*,*84*(2), 327-332.
- Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *cognition*, *58*(1), 1-73.
- Dewey, J. (1933). *How we think*. New York: Health and Company.
- Dewey, J. (1997). Experience and education. 1938. *New York: First Touchstone Edition*, 64-67.
- Elder, L., & Paul, R. (2010). Critical Thinking: Competency Standards Essential for the Cultivation of Intellectual Skills, Part 1. *Journal of Developmental Education*, *34*(2), 38-39.
- Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. Educational Researcher, 18(3), 4–10
- Ennis, R. H. (1991). Critical thinking: A streamlined conception. *Teaching philosophy*, *14*(1), 5-24.
- Ennis, R. H. (2011). The nature of critical thinking: An outline of critical thinking dispositions and abilities. *University of Illinois*, 2-4.
- Ennis, R. H. (2016). Definition: A Three-Dimensional Analysis with Bearing on Key Concepts. In Proceedings of the Ontario Society for the Study of Argumentation Conference (Vol. 11, p. 150).
- Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on psychological science*, *8*(3), 223-241.
- Facione, P. A. (1990). Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction. Research Findings and Recommendations.
- Facione, P. A. (2011). Critical thinking: What it is and why it counts. *Insight assessment*, *2007*(1), 1-23.
- Foundation for critical thinking, https://www.criticalthinking.org/pages/defining-critical-thinking/766
- Gelder, T. V. (2005). Teaching critical thinking: Some lessons from cognitive science. College teaching, 53(1), 41-48.
- Gentner, D. (1989). Analogical learning. *Similarity and analogical reasoning*, *199*.

- Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond "heuristics and biases". *European review of social psychology*, *2*(1), 83-115.
- Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky.
- Gigerenzer, G. (2004). Fast and frugal heuristics: The tools of bounded rationality. *Blackwell handbook of judgment and decision making*, *62*, 88.
- Gigerenzer, G., Hoffrage, U., & Goldstein, D. G. (2008). *Fast and frugal heuristics are plausible models of cognition: Reply to Dougherty*, Franco-Watkins, and Thomas (2008).
- Gigerenzer, G., & Todd, P. M. (1999). ABC Research Group. *Simple heuristics that make us smart*.
- Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge university press.
- Glaser, E. M. (1941). *An experiment in the development of critical thinking*. Teachers College, Columbia University.
- Gopnik, A., Schulz, L., & Schulz, L. E. (Eds.). (2007). *Causal learning: Psychology, philosophy, and computation*. Oxford University Press.
- Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Disposition, skills, structure training, and metacognitive monitoring. *American psychologist*, *53*(4), 449.
- Halpern, D. F. (1999). Teaching for critical thinking: Helping college students develop the skills and dispositions of a critical thinker. *New directions for teaching and learning*, *1999*(80), 69-74.
- Halpern, D. F. (2007). The nature and nurture of critical thinking. *Critical thinking in psychology*, 1-14.
- Halpern, D. F. (2013). *Thought and knowledge: An introduction to critical thinking*. Psychology Press.
- Haselton, M. G., Bryant, G. A., Wilke, A., Frederick, D. A., Galperin, A., Frankenhuis, W. E., & Moore, T. (2009). Adaptive rationality: An evolutionary perspective on cognitive bias. *Social Cognition*, *27*(5), 733-763.
- Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and social psychology Review*, *10*(1), 47-66.
- Hitchcock, D. (2018). Critical thinking. Metaphysics Research Lab, Stanford University. Retrieved from https://plato.stanford.edu/archives/fall2018/entries/critical-thinking/
- Kahan, D. M. (2015). The politically motivated reasoning paradigm, part 1: What politically motivated reasoning is and how to measure it. *Emerging trends in the social and behavioral sciences: An interdisciplinary, searchable, and linkable resource*, 1-16.
- Kahan, D. M., Jenkins Smith, H., & Braman, D. (2011). Cultural cognition of scientific consensus. *Journal of risk research*, *14*(2), 147-174.
- Kahan, D. M., Braman, D., Cohen, G. L., Gastil, J., & Slovic, P. (2010). Who fears the HPV vaccine, who doesn't, and why? An experimental study of the mechanisms of cultural cognition. *Law and human behavior*, *34*(6), 501-516.
- Kahan, D. M., & Stanovich, K. (2016). Rationality and belief in human evolution. *Annenberg Public Policy Center Working Paper*, (5)
- Kahneman, D., & Tversky, A. (1996). *On the reality of cognitive illusions*.

- Kahneman, D., Slovic, S. P., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge university press.
- Kuhn, D. (1992). Thinking as argument. *Harvard Educational Review*, *62*(2), 155-179.
- Kuhn, D. (1999). A developmental model of critical thinking. *Educational researcher*, *28*(2), 16-46.
- Kuhn, D. (2018). A role for reasoning in a dialogic approach to critical thinking. Topoi, 37(1), 121-128.
- Lai, E. R. (2011). Critical thinking: A literature review. *Pearson's Research Reports*, *6*, 40-41.
- Lewis, A., & Smith, D. (1993). Defining Higher Order Thinking. *Theory into Practice*, 32(3,), 131–137.
- Lipman, M. (2003). Thinking in education. (2nd. ed.). New York: Cambridge University Press.
- Lipman, M.. (1987). Critical thinking: What can it be? *Analytic Teaching*, 8(1).
- Lipman, M. (1988). Critical thinking: What can it be? Resource Publication, Series 1 No. 1.
- McPeck, J. E. (1990). Critical thinking and subject specificity: A reply to Ennis. *Educational researcher*, *19*(4), 10-12.
- Mercier, H. (2017). How gullible are we? A review of the evidence from psychology and social science. *Review of General Psychology*, *21*(2), 103-122.
- Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory.
- Mercier, H., & Sperber, D. (2017). *The enigma of reason*. Harvard University Press.
- Nisbett, R. E. (2015). *Mindware: Tools for smart thinking*. Farrar, Straus and Giroux.
- Nisbett, R. E., Fong, G. T., Lehman, D. R., & Cheng, P. W. (1987). Teaching reasoning. *Science*, *238*(4827), 625-631.
- Oreskes, N. (2019). *Why Trust Science?* (Vol. 1). Princeton University Press.
- Origgi, G. (2019). *Reputation: What it is and why it Matters*. Princeton University Press.
- Paul, R. & Elder. L. (2009). *The Miniature Guide to Critical Thinking Concepts and Tools*, Foundation for Critical Thinking, Dillon Beach, CA, 5th Ed.
- Paul, R. W., Elder, L., & Bartell, T. (1997). *California teacher preparation for instruction in critical thinking: Research findings and policy recommendations*.
- Siegel, H. (2010). Critical thinking. In *International encyclopedia of education* (pp. 141-145). Elsevier Ltd.
- Scriven, M., & Paul, R. (1987). Critical thinking as defined by the National Council for Excellence in Critical Thinking. In *8th Annual International Conference on Critical Thinking and Education Reform, Rohnert Park, CA* (pp. 25-30).
- Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental science*, *10*(1), 89-96.
- Stanovich, K. E., & Stanovich, P. J. (2010). A framework for critical thinking, rational thinking, and intelligence. *Innovations in educational psychology: Perspectives on learning, teaching and human development*, 195-237.
- Sternberg, R. J. (1986). Critical Thinking: Its Nature, Measurement, and Improvement.
- Trouche, E., Johansson, P., Hall, L., & Mercier, H. (2016). The selective laziness of reasoning. *Cognitive Science*, *40*(8), 2122-2136.

- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, *185*(4157), 1124-1131.
- West, R. F., Toplak, M. E., & Stanovich, K. E. (2008). Heuristics and biases as measures of critical thinking: Associations with cognitive ability and thinking dispositions. *Journal of Educational Psychology*, *100*(4), 930.
- Willingham, D. T. (2007). Critical thinking: Why it is so hard to teach? *American federation of teachers summer 2007, pp. 8-19*.
- Willingham, D. T. (2019). How to teach critical thinking. Education: Future Frontiers Occasional Paper Series. Retrieved from Analysis and Policy Observatory Website: https://apo.org.au/node/244676

Lectures supplémentaires utilisées

Sharon Bailin

- Bailin, S. (1987). Critical and creative thinking. *Informal logic*, 9(1).
- Bailin, S., Battersby, M., & Clauss, P. (2011). *Reason in the balance: Teaching critical thinking as dialectical*. 13.
- Bailin, S., Case, R., Coombs, J. R., & Daniels, L. B. (1999a). Common misconceptions of critical thinking. *Journal of Curriculum Studies*, 31(3), 269–283.
- Bailin, S., & Siegel, H. (2003). Critical thinking. The Blackwell Guide to the *Philosophy of Education*, 181, 193.
- Bailin, S. (1998). *Skills, generalizibility and critical thinking. In twentieth world congress on philosophy*. Boston: The Paideia Archive.
- Battersby, M., & Bailin, S. (2011). Critical Inquiry: Considering the Context. *Argumentation*, 25(2), 243–253.

Robert Ennis

- Ennis, R. H. (1962). A concept of critical thinking. *Harvard educational review*.
- Ennis, R. H. (1964). A definition of critical thinking. *The Reading Teacher*, *17*(8), 599-612.
- Ennis, R. H. (1993). Critical thinking assessment. *Theory into practice*, *32*(3), 179-186.
- Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. *Educational researcher*, *18*(3), 4-10.
- Ennis, R. H. (1984). Problems in testing informal logic critical thinking reasoning ability. *Informal Logic*, *6*(1).
- Ennis, R. H. (1985). A logical basis for measuring critical thinking skills. *Educational leadership*, *43*(2), 44-48.
- Ennis, R. (1991). Critical thinking: A streamlined conception. *Teaching philosophy*, *14*(1), 5-24.
- Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities.
- Ennis, R. H. (1996). Critical thinking dispositions: Their nature and assessability. *Informal logic*, *18*(2).
- Ennis, R. H. (1987). Critical thinking and the curriculum. *Thinking skills instruction: Concepts and techniques*, 40-48.
- Kennedy, M., Fisher, M. B., & Ennis, R. H. (1991). Critical thinking: Literature review and needed research. *Educational values and cognitive instruction: Implications for reform*, *2*, 11-40.

- Norris, S. P., & Ennis, R. H. (1989). *Evaluating Critical Thinking. The Practitioners' Guide to Teaching Thinking Series*. Critical Thinking Press and Software, Box 448, Pacific Grove, CA 93950-0448; tele.
- Ennis, R. H. (1990). The extent to which critical thinking is subject-specific: Further clarification. *Educational researcher*, *19*(4), 13-16.
- Ennis, R. (2011). Critical thinking: Reflection and perspective Part II. *Inquiry: Critical thinking across the Disciplines*, *26*(2), 5-19.
- Ennis, R. H. (1998). Is critical thinking culturally biased? *Teaching philosophy*, *21*(1), 15-33.
- Ennis, R. H. (2015). Critical thinking: A streamlined conception. In *The Palgrave handbook of critical thinking in higher education* (pp. 31-47). Palgrave Macmillan, New York.
- Ennis, R. (1997). Incorporating critical thinking in the curriculum: An introduction to some basic issues. *Inquiry: Critical Thinking across the disciplines*, *16*(3), 1-9.

Peter Facione

- Facione, P. A., Sanchez, C. A., Facione, N. C., & Gainen, J. (1995). The disposition toward critical thinking. *The Journal of General Education*, 1-25.
- Facione, N. C., Facione, P. A., & Sanchez, C. A. (1994). Critical thinking disposition as a measure of competent clinical judgment: The development of the California Critical Thinking Disposition Inventory. *Journal of Nursing Education*, *33*(8), 345-350.
- Facione, P. A. (2000). The disposition toward critical thinking: Its character, measurement, and relationship to critical thinking skill. *Informal logic*, *20*(1).
- Facione, P. A., & Facione, N. C. (1994). Holistic critical thinking scoring rubric. *Retrieved July*, *19*, 2007.
- Giancarlo, C. A., & Facione, P. A. (2001). A look across four years at the disposition toward critical thinking among undergraduate students. *The Journal of General Education*, *50*(1), 29-55.
- Facione, N. C., & Facione, P. A. (2008). Critical thinking and clinical judgment. *Critical thinking and clinical reasoning in the health sciences: A teaching anthology*, 1-13.
- Facione, P. A., Facione, N. C., & Giancarlo, C. A. (1997). Professional judgment and the disposition toward critical thinking. *Retrieved August*, *1*, 2004.
- Facione, P. A. (1990). The California Critical Thinking Skills Test--College Level. Technical Report# 2. Factors Predictive of CT Skills.
- Facione, P. A. (1990). The California Critical Thinking Skills Test--College Level. Technical Report# 1. Experimental Validation and Content Validity.
- Facione, P. A. (1984). Toward a theory of critical thinking. *Liberal Education*, *70*(3), 253-61.
- Facione, P. A. (1991). Using the California Critical Thinking Skills Test in Research, Evaluation, and Assessment.
- Facione, P. A., & Facione, N. C. (2007). Talking critical thinking. *Change: The Magazine of Higher Learning*, *39*(2), 38-45.

Diane Halpern / related to Halpern Critical Thinking Assessment

- Halpern, D. F. (1993). Assessing the effectiveness of critical-thinking instruction. *The Journal of General Education*, 238-254.

- Halpern, D. F. (2014). *Critical thinking across the curriculum: A brief edition of thought & knowledge*. Routledge.
- Halpern, D. F. (1987). Analogies as a critical thinking skill. *Applications of cognitive psychology: Problem solving, education, and computing*, 75-86.
- Halpern, D. F. (2006). Is intelligence critical thinking? Why we need a new definition of intelligence. *Extending intelligence: Enhancement and new constructs*, *293310*.
- HAU, K. T., HALPERN, D. F., MARIN-BURKHART, L., HO, T., Ku, Y. L., CHAN, N. M., & LUN, M. C. (2006). Chinese and United States Students' critical thinking: Cross-cultural construct validation of a critical thinking assessment.
- Butler, H. A. (2012). Halpern Critical Thinking Assessment predicts real world outcomes of critical thinking. *Applied Cognitive Psychology*, *26*(5), 721-729.
- Marin, L. M., & Halpern, D. F. (2011). Pedagogy for developing critical thinking in adolescents: Explicit instruction produces greatest gains. *Thinking Skills and Creativity*, *6*(1), 1-13.
- Butler, H. A., Dwyer, C. P., Hogan, M. J., Franco, A., Rivas, S. F., Saiz, C., & Almeida, L. S. (2012). The Halpern Critical Thinking Assessment and real-world outcomes: Cross-national applications. *Thinking Skills and Creativity*, *7*(2), 112-121.
- Sternberg, R. J., Roediger III, H. L., & Halpern, D. F. (Eds.). (2007). *Critical thinking in psychology*. Cambridge University Press.

Deanna Kuhn

- Dean, D., & Kuhn, D. (2003). Metacognition and Critical Thinking.
- Kuhn, D., & Dean, Jr, D. (2004). Metacognition: A bridge between cognitive psychology and educational practice. *Theory into practice*, *43*(4), 268-273.
- Kuhn, D. (1993). Science as argument: Implications for teaching and learning scientific thinking. *Science education*, *77*(3), 319-337.
- Kuhn, D. (2015). Thinking together and alone. *Educational researcher*, *44*(1), 46-53.
- Felton, M. K., & Kuhn, D. (2007). "How do I Know?" The Epistemological Roots of Critical Thinking. *Journal of Museum Education*, *32*(2), 101-110.
- Kuhn, D. (1996). Is good thinking scientific thinking. *Modes of thought: Explorations in culture and cognition*, 261-281.
- Kuhn, D. (2018). A role for reasoning in a dialogic approach to critical thinking. *Topoi*, *37*(1), 121-128
- Kuhn, D., & Crowell, A. (2011). Dialogic argumentation as a vehicle for developing young adolescents' thinking. *Psychological science*, *22*(4), 545-552.
- Kuhn, D. (2005). *Education for thinking*. Harvard University Press.
- Kuhn, D., Hemberger, L., & Khait, V. (2017). *Argue with me: Argument as a path to developing students' thinking and writing*. Routledge.
- *Kuhn, D. (2000). Metacognitive development. Current directions in psychological science, 9(5), 178-181.*
- *Kuhn, D., Cheney, R., & Weinstock, M. (2000). The development of epistemological understanding. Cognitive development, 15(3), 309-328.*
- *Kuhn, D. (1991). Education for thinking: What can psychology contribute? Promoting cognitive growth over the life span.*
- *Kuhn, D. (1992). Thinking as argument. Harvard Educational Review, 62(2), 155-179.*
- *Kuhn, D., & Udell, W. (2003). The development of argument skills. Child development, 74(5), 1245-1260.*

Matthew Lipman

- Lipman, M. (2003). *Thinking in education*. Cambridge University Press.
- Lipman, M. (1974). *Harry Stottlemeier's discovery*.
- Lipman, M. (1985). Thinking skills fostered by philosophy for children. *Thinking and Learning Skill Vol. 1: Relating Instruction to Research*, 83-108.
- Lipman, M. (1984). The Cultivation of Reasoning through Philosophy. *Educational leadership*, *42*(1), 51-56.
- Lipman, M. (1998). Teaching students to think reasonably: Some findings of the Philosophy for Children program. *The Clearing House*, *71*(5), 277-280.
- Lipman, M. (1987). *Some thoughts on the foundations of reflective education*.
- Lipman, M. (1993). Promoting better classroom thinking. *Educational Psychology*, *13*(3-4), 291-304.
- Lipman, M. (1996). Philosophical discussion plans and exercises. *Analytic Teaching*, *16*(2).
- Lipman, M., & Sharp, A. M. (1978). Some educational presuppositions of philosophy for children. *Oxford review of Education*, *4*(1), 85-90.
- Lipman, M. (1988). Philosophy for children and critical thinking. *Thinking: The Journal of Philosophy for Children*, *7*(4), 40-42.
- Lipman, M. (1997). Thinking in community. *Inquiry: Critical thinking across the disciplines*, *16*(4), 6-21.
- Lipman, M. (1985). Philosophy and the Cultivation of Reasoning. *Thinking: The Journal of Philosophy for Children*, *5*(4), 33-41.
- Lipman, M. (1995). Good thinking. *Inquiry: Critical Thinking Across the Disciplines*, *15*(2), 37-41.
- Lipman, M. (1988). Critical thinking and the use of criteria. *Inquiry: Critical Thinking across the Disciplines*, *1*(2), 2-2.
- Lipman, M. (2011). Philosophy for children: Some assumptions and implications. *Ethics in Progress*, *2*(1), 3-16.
- Lipman, M. (1954). The relation of critical functions and critical decisions to art inquiry. *The Journal of Philosophy*, *51*(22), 653-667.
- Lipman, M. (1987). Preparing teachers to teach for thinking. *Philosophy Today*, *31*(1), 90-96.
- Lipman, M. (1992). Unreasonable people and inappropriate judgments. *Inquiry: Critical Thinking Across the Disciplines*, *10*(3), 1-1.
- Lipman, M. (1992). Criteria and judgment in critical thinking. *Inquiry: Critical Thinking Across the Disciplines*, *9*(4), 3-4.
- Lipman, M. (1982). Why Aren't Thinking Skills Being Taught? *Thinking: The Journal of Philosophy for Children*, *3*(3/4), 45-46.

Richard Nisbett

- Lehman, D. R., & Nisbett, R. E. (1990). A longitudinal study of the effects of undergraduate training on reasoning. *Developmental Psychology*, *26*(6), 952.
- Norenzayan, A., Smith, E. E., Kim, B. J., & Nisbett, R. E. (2002). Cultural preferences for formal versus intuitive reasoning. *Cognitive science*, *26*(5), 653-684.
- Nisbett, R. E. (2013). *Rules for reasoning*. Psychology Press.

- Fong, G. T., & Nisbett, R. E. (1991). Immediate and delayed transfer of training effects in statistical reasoning. *Journal of Experimental Psychology: General*, *120*(1), 34.
- Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive psychology*, 18(3), 253-292.
- E. Nisbett, R., & Norenzayan, A. (2002). *Culture and cognition. Stevens' handbook of experimental psychology*.
- Nisbett, R. E., Fong, G. T., Lehman, D. R., & Cheng, P. W. (1987). Teaching reasoning. *Science*, 238(4827), 625-631.

Richard Paul, Linda Elder

- Elder, L., & Paul, R. (2012). Critical Thinking: Competency Standards Essential to the Cultivation of Intellectual Skills, Part 4. *Journal of Developmental Education*, 35(3), 30–31.
- Elder, L., & Paul, R. (1994). Critical thinking: Why we must transform our teaching. *Journal of Developmental Education*, 18(1), 34.
- Elder, L., & Paul, R. (2008). Critical thinking: Strategies for improving student learning. *Journal of Developmental Education*, 32(1), 32.
- Elder, L., & Paul, R. (2009). *The aspiring thinker's guide to critical thinking*.
- Elder, L., & Paul, R. (2019). *The thinker's guide to analytic thinking: How to take thinking apart and what to look for when you do*. Rowman & Littlefield.
- Paul, R. (1981). Teaching critical thinking in the "strong" sense: A focus on self-deception, world views, and a dialectical mode of analysis. *Informal Logic*, 4(2).
- Paul, R. W. (1984). Critical thinking: fundamental to education for a free society. *Educational leadership*, 42(1), 4-14.
- Paul, R. W. (1985). Bloom's Taxonomy and Critical Thinking Instruction. *Educational leadership*, 42(8), 36-39.
- Paul, R. W. (1985). The critical-thinking movement. In *National Forum* (Vol. 65, No. 1, p. 2). Honor Society of Phi Kappa Phi.
- Paul, R. (1989). *Critical Thinking Handbook: High School. A Guide for Redesigning Instruction*. Center for Critical Thinking and Moral Critique, Sonoma State University, Rohnert Park, CA 94928.
- Paul, R. W., & Binker, A. J. A. (1990). *Critical thinking: What every person needs to survive in a rapidly changing world*. Center for Critical Thinking and Moral Critique, Sonoma State University, Rohnert Park, CA 94928.
- Paul, R. (1990). Critical thinking. Rohnert Park, CA: Sonoma State University.
- Paul, R., & Nosich, G. M. (1992). *A Model for the National Assessment of Higher Order Thinking*.
- Paul, R. W. (1993). The logic of creative and critical thinking. *American Behavioral Scientist*, 37(1), 21-39.
- Elder, L., & Paul, R. (1997). Critical thinking: Crucial distinctions for questioning. *Journal of Developmental Education*, *21*(2), 34.
- Paul, R. (2000). *Critical thinking, moral integrity and citizenship: Teaching for the intellectual virtues. Knowledge, belief and character: Readings in virtue epistemology*, 163-175.
- Paul, R., & Elder, L. (2003). Critical thinking: Teaching students how to study and learn (Part III). *Journal of Developmental Education*, 26(3), 36.

- Paul, R. (2005). *The state of critical thinking today. New directions for community colleges*, 2005(130), 27-38.
- Paul, R., & Elder, L. (2006). Critical thinking: The nature of critical and creative thought. *Journal of Developmental Education*, 30(2), 34.
- Paul, R., & Elder, L. (2007). Critical thinking: The art of Socratic questioning. *Journal of developmental education*, 31(1), 36.
- Paul, R., & Elder, L. (2013). *Critical thinking: Tools for taking charge of your professional and personal life*. Pearson Education.
- Paul, R. W. (2018). Critical thinking and the critical person. In *Thinking* (pp. 373-403). Routledge.
- Paul, R., & Elder, L. (2019). *A guide for educators to critical thinking competency standards: Standards, principles, performance indicators, and outcomes with a critical thinking master rubric*. Rowman & Littlefield.
- Paul, R., & Elder, L. (2019). *The nature and functions of critical & creative thinking*. Rowman & Littlefield.
- Walsh, D., & Paul, R. W. (1986). *The Goal of Critical Thinking: from Educational Ideal to Educational Reality*.

Keith Stanovich

- Stanovich, K. E., Sá, W. C., & West, R. F. (2004). Individual differences in thinking, reasoning, and decision making. *The nature of reasoning*, 375-409.
- Stanovich, K. E., Toplak, M. E., & West, R. F. (2008). The development of rational thought: A taxonomy of heuristics and biases. In *Advances in child development and behavior* (Vol. 36, pp. 251-285). JAI.
- Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & cognition*, *39*(7), 1275.
- Sá, W. C., West, R. F., & Stanovich, K. E. (1999). The domain specificity and generality of belief bias: Searching for a generalizable critical thinking skill. *Journal of educational psychology*, *91*(3), 497.
- Kokis, J. V., Macpherson, R., Toplak, M. E., West, R. F., & Stanovich, K. E. (2002). Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles. *Journal of Experimental Child Psychology*, *83*(1), 26-52.

Tim van Gelder

- Van Gelder, T. (2001, December). How to improve critical thinking using educational technology. In *Meeting at the crossroads: Proceedings of the 18th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education* (pp. 539-548).

Autres

- Byrnes, J. P., & Dunbar, K. N. (2014). The nature and development of critical-analytic thinking. *Educational Psychology Review*, *26*(4), 477-493.
- Dwyer, C. P., Hogan, M. J., & Stewart, I. (2014). An integrated critical thinking framework for the 21st century. *Thinking Skills and Creativity*, *12*, 43-52.
- Hitchcock, D. (2017). Do the fallacies have a place in the teaching of reasoning skills or critical thinking? In *On reasoning and argument* (pp. 401-408). Springer, Cham

- Lilienfeld, S. O., Ammirati, R., & Landfield, K. (2009). Giving debiasing away: Can psychological research on correcting cognitive errors promote human welfare? *Perspectives on psychological science*, *4*(4), 390-398.
- McPeck, J. E. (2016). *Critical thinking and education*. Routledge.
- Siegel, H. (1991). The generalizability of critical thinking. *Educational Philosophy and Theory*, *23*(1), 18-30.
- Siegel, H. (1980, November). Critical thinking as an educational ideal. In *The Educational Forum* (Vol. 45, No. 1, pp. 7-23). Taylor & Francis Group.
- Siegel, H. (2013). *Educating reason*. Routledge.

Général

- Baron, J. (1993). Why Teach Thinking? An Essay. *Applied Psychology*, 42(3), 191–214.
- Boisvert, J. (2015). Pensée critique: définition, illustration et applications. *Revue Québécoise de Psychologie*, Volume 36 Numéro 1, Pages 3-33.
- Pithers, R. T., & Soden, R. (2000). Critical thinking in education: A review. *Educational research*, *42*(3), 237-249.

3. Bases cognitives

- Acerbi, A. (2019a). Cognitive attraction and online misinformation. *Palgrave Communications*, *5*(1), 1-7.
- Acerbi, A. (2019b). *Cultural Evolution in the Digital Age*. Oxford University Press.
- Allom, V., Mullan, B., & Hagger, M. (2016). Does inhibitory control training improve health behaviour? A meta-analysis. *Health Psychology Review*, *10*(2), 168-186.
- Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states?. *Psychological review*, *116*(4), 953.
- Ariely, D. (2008). *Predictably irrational*. New York, NY: Harper Audio.
- Aron, A. R. (2007). The neural basis of inhibition in cognitive control. *The neuroscientist*, *13*(3), 214-228.
- Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: one decade on. *Trends in cognitive sciences*, *18*(4), 177-185.
- Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. *Organizational influence processes*, 295-303.
- Atran, S. (2002). *In gods we trust: The evolutionary landscape of religion*. Oxford University Press.
- Bago, B., & De Neys, W. (2017). Fast logic?: Examining the time course assumption of dual process theory. *Cognition*, *158*, 90-109.
- Bago, B., & De Neys, W. (2019). The smart System 1: Evidence for the intuitive nature of correct responding on the bat-and-ball problem. *Thinking & Reasoning*, *25*(3), 257-299.
- Ball, LJ, Thompson, VA, & Stupple, EJN (2017). Conflict and dual process theory: The case of belief bias. In W. De Neys (Ed.), Dual Process Theory 2.0 (pp. 100-120). Abingdon, Oxon: Routledge. Conflict and Dual-Process Theory: The Case of Belief Bias.
- Battersby, M., & Bailin, S. (2013). Critical thinking and cognitive biases.

- Bebbington, K., MacLeod, C., Ellison, T. M., & Fay, N. (2017). The sky is falling: evidence of a negativity bias in the social transmission of information. *Evolution and Human Behavior*, *38*(1), 92-101.
- Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, *275*(5304), 1293-1295.
- Białek, M., & De Neys, W. (2017). Dual processes and moral conflict: Evidence for deontological reasoners' intuitive utilitarian sensitivity. *Judgment and Decision Making*, *12*(2), 148-167.
- Blaine, T., & Boyer, P. (2018). Origins of sinister rumors: A preference for threat-related material in the supply and demand of information. *Evolution and Human Behavior*, *39*(1), 67-75.
- Boyd, R., & Richerson, P. J. (2005). *The origin and evolution of cultures*. Oxford University Press.
- Boyer, P. (2018). *Minds make societies: How cognition explains the world humans create*. Yale University Press.
- Butler, A. C., Karpicke, J. D., & Roediger III, H. L. (2008). Correcting a metacognitive error: feedback increases retention of low-confidence correct responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 918.
- Butterfield, B., & Metcalfe, J. (2006). The correction of errors committed with high confidence. *Metacognition and Learning*, *1*(1), 69-84.
- Carpenter, J., Sherman, M. T., Kievit, R. A., Seth, A. K., Lau, H., & Fleming, S. M. (2019). Domain-general enhancements of metacognitive ability through adaptive training. *Journal of Experimental Psychology: General*, *148*(1), 51.
- Casscells, W., Schoenberger, A., & Graboys, T. B. (1978). Interpretation by physicians of clinical laboratory results. *New England Journal of Medicine*, *299*(18), 999-1001.
- Cavalli-Sforza, L. L., & Feldman, M. W. (1981). *Cultural transmission and evolution: A quantitative approach* (No. 16). Princeton University Press.
- Chi, M., Glaser, R., & Rees, E. (1982). Experience and problem solving. *Advances in the psychology of human intelligence*, *1*.
- Chudek, M., & Henrich, J. (2011). Culture–gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in cognitive sciences*, *15*(5), 218-226.
- Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, *31*(3), 187-276.
- Cosmides, L., & Tooby, J. (1994). Better than rational : Evolutionary psychology and the invisible hand. *The American Economic Review*,*84*(2), 327-332.
- Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *cognition*, *58*(1), 1-73.
- De Neys, W., Vartanian, O., & Goel, V. (2008). Smarter than we think: When our brains detect that we are biased. *Psychological Science*, *19*(5), 483-489.
- De Neys, W., Moyens, E., & Vansteenwegen, D. (2010). Feeling we're biased: Autonomic arousal and reasoning conflict. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(2), 208-216.
- De Neys, W., Rossi, S., & Houdé, O. (2013). Bats, balls, and substitution sensitivity: Cognitive misers are no happy fools. *Psychonomic Bulletin & Review*, *20*(2), 269-273.

- De Neys, W., Lubin, A., & Houdé, O. (2014). The smart nonconserver: Preschoolers detect their number conservation errors. *Child Development Research*, *2014*.
- Dawkins, R. (1982). Replicators and vehicles. *Current problems in sociobiology*, *45*, 64.
- Dawkins, R. (2016). *The selfish gene*. Oxford university press.
- Diamond, A. (1991). Neuropsychological insights into the meaning of object concept development.
- Diamond, A., & Ling, D. S. (2016). Conclusions about interventions, programs, and approaches for improving executive functions that appear justified and those that, despite much hype, do not. *Developmental cognitive neuroscience*, *18*, 34-48.
- Dunlosky, J., & Metcalfe, J. (2008). *Metacognition*. Sage Publications.
- Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current directions in psychological science*, *12*(3), 83-87.
- Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist*, *13*(4), 277-287.
- Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational psychologist*, *46*(1), 6-25.
- Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., & Kruger, J. (2008). Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent. *Organizational behavior and human decision processes*, *105*(1), 98-121.
- Enge, S., Behnke, A., Fleischhauer, M., Küttler, L., Kliegel, M., & Strobel, A. (2014). No evidence for true training and transfer effects after inhibitory control training in young healthy adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(4), 987.
- Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on psychological science*, *8*(3), 223-241.
- Fischhoff, B. (1982). Debiasing. Judgment under uncertainty: Heuristics and biases. *Judgment under uncertainty: Heuristics and biases*, 422-444.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American psychologist*, *34*(10), 906.
- Fleming, S. M., & Daw, N. D. (2017). Self-evaluation of decision-making: A general Bayesian framework for metacognitive computation. Psychological Review, 124(1), 91.
- Fleming, S. M., Dolan, R. J., & Frith, C. D. (2012). Metacognition: computation, biology and function. The Royal Society.
- Fleming, S. M., & Frith, C. D. (Eds.). (2014). *The cognitive neuroscience of metacognition*. London, UK: Springer.
- Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. Frontiers in Human Neuroscience, 8, 443.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives*, *19*(4), 25-42.
- Galef Jr, B. G., Dudley, K. E., & Whiskin, E. E. (2008). Social learning of food preferences in 'dissatisfied'and 'uncertain'Norway rats. *Animal Behaviour*, *75*(2), 631-637.
- Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond "heuristics and biases". *European review of social psychology*, *2*(1), 83-115.

- Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky.
- Gigerenzer, G. (2004). Fast and frugal heuristics: The tools of bounded rationality. *Blackwell handbook of judgment and decision making*, *62*, 88.
- Gigerenzer, G., Hoffrage, U., & Goldstein, D. G. (2008). *Fast and frugal heuristics are plausible models of cognition: Reply to Dougherty*, Franco-Watkins, and Thomas (2008).
- Gigerenzer, G., & Hug, K. (1992). Domain-specific reasoning: Social contracts, cheating, and perspective change. *Cognition*, *43*(2), 127-171.
- Gigerenzer, G., & Todd, P. M. the ABC Research Group 1999. *Simple heuristics that make us smart*.
- Gilovich, T., & Griffin, D. (2002). Introduction-heuristics and biases: Then and now. *Heuristics and biases: The psychology of intuitive judgment*, 1-18.
- Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge university press.
- Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms in very young children: Two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental psychology*, *37*(5), 620.
- Goupil, L., & Kouider, S. (2016). Behavioral and neural indices of metacognitive sensitivity in preverbal infants. Current Biology, 26(22), 3038–3045.
- Goupil, L., & Kouider, S. (2019). Developing a Reflective Mind: From Core Metacognition to Explicit Self-Reflection. Current Directions in Psychological Science, 0963721419848672.
- Goupil, L., Romand-Monnier, M., & Kouider, S. (2016). Infants ask for help when they know they don't know. *Proceedings of the National Academy of Sciences*, *113*(13), 3492-3496.
- Grimaldi, P., Lau, H., & Basso, M. A. (2015). There are things that we know that we know, and there are things that we do not know we do not know: Confidence in decision-making. Neuroscience & Biobehavioral Reviews, 55, 88–97.
- Halpern, D. F. (2013). *Thought and knowledge: An introduction to critical thinking*. Psychology Press.
- Harris, P. L. (2012). Trusting what you're told: How children learn from others. Harvard University Press.
- Harris, P. L., & Corriveau, K. H. (2011). Young children's selective trust in informants. Philosophical Transactions of the Royal Society B: Biological Sciences, 366(1567), 1179–1187.
- Haselton, M. G., Bryant, G. A., Wilke, A., Frederick, D. A., Galperin, A., Frankenhuis, W. E., & Moore, T. (2009). Adaptive rationality: An evolutionary perspective on cognitive bias. *Social Cognition*, *27*(5), 733-763.
- Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and social psychology Review*, *10*(1), 47-66.
- Haselton, M. G., Nettle, D., & Murray, D. R. (2015). The evolution of cognitive bias. *The handbook of evolutionary psychology*, 1-20.
- Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of educational research*, *66*(2), 99-136.

- Heath, C., Bell, C., & Sternberg, E. (2001). Emotional selection in memes: the case of urban legends. *Journal of personality and social psychology*, *81*(6), 1028.
- Heilbron, M., & Meyniel, F. (2019). Confidence resets reveal hierarchical adaptive learning in humans. *PLoS computational biology*, *15*(4), e1006972.
- Henrich, J., & Gil-White, F. J. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and human behavior*, *22*(3), 165-196.
- Higgins, S., Katsipataki, M., Villanueva-Aguilera, A. B., Coleman, R., Henderson, P., Major, L. E., ... & Mason, D. (2016). The Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit.
- Houdé, O., & Borst, G. (2015). Evidence for an inhibitory-control theory of the reasoning brain. *Frontiers in human neuroscience*, *9*, 148.
- Hutchins, E. (1995). *Cognition in the Wild* (No. 1995). MIT press.
- Hutchins, E. (1995). How a cockpit remembers its speeds. *Cognitive science*, *19*(3), 265-288.
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
- Kahneman, D., Slovic, S. P., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge university press.
- Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions.
- Kepecs, A., & Mainen, Z. F. (2012). A computational framework for the study of confidence in humans and animals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1322-1337.
- Kepecs, A., Uchida, N., Zariwala, H. A., & Mainen, Z. F. (2008). Neural correlates, computation and behavioural impact of decision confidence. *Nature*, *455*(7210), 227-231.
- Klahr, D., Fay, A. L., & Dunbar, K. (1993). Heuristics for scientific experimentation: A developmental study. *Cognitive psychology*, *25*(1), 111-146.
- Knill, D. C., & Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *TRENDS in Neurosciences*, *27*(12), 712-719.
- Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological review*, *100*(4), 609.
- Koriat, A. (2012). The self-consistency model of subjective confidence. *Psychological review*, *119*(1), 80.
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of personality and social psychology*, *77*(6), 1121.
- Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological review*, *96*(4), 674.
- Kuhn, D. (2005). *Education for thinking*. Harvard University Press.
- Kuhn, D. (2011). What is scientific thinking and how does it develop?.
- Kuhn, D., Garcia-Mila, M., Zohar, A., Andersen, C., White, S. H., Klahr, D., & Carver, S. M. (1995). Strategies of knowledge acquisition. *Monographs of the society for research in child development*, i-157.
- Jiménez, Á. V., & Mesoudi, A. (2019). Prestige-biased social learning: current evidence and outstanding questions. *Palgrave Communications*, *5*(1), 1-12.
- Lea, A. M., & Ryan, M. J. (2015). Irrationality in mate choice revealed by túngara frogs. *Science*, *349*(6251), 964-966.

- Lhermitte, F. (1983). 'Utilization behaviour'and its relation to lesions of the frontal lobes. *Brain*, *106*(2), 237-255.
- Lichtenstein, S. B. Fischoff, and L. Phillips (1982),"Calibration of Probabilities: The State of the Art to 1980,". *Judgement under Uncertainty: Heuristics and Biases*, 306-334.
- Lsockl, K., & Schneider, W. (2007). Knowledge about the mind: Links between theory of mind and later metamemory. *Child Development*, *78*(1), 148-167.
- Mascaro, O., & Morin, O. (2015). Epistemology for beginners: Two-to five-year-old children's representation of falsity. *PloS one*, *10*(10).
- Mercier, H. (2017). How gullible are we? A review of the evidence from psychology and social science. *Review of General Psychology*, *21*(2), 103-122.
- Mercier, H. (2020). *Not born yesterday: The science of who we trust and what we believe*. Princeton University Press.
- Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory.
- Metcalfe, J., & Shimamura, A. P. (Eds.). (1994). *Metacognition: Knowing about knowing*. MIT press.
- Meyniel, F., Sigman, M., & Mainen, Z. F. (2015). Confidence as Bayesian probability: From neural origins to behavior. *Neuron*, *88*(1), 78-92
- Meyniel, F. A bayesian approach to the brain. dana.org/article/a-bayesian-approach-to-the-brain/
- Meyniel, F., & Dehaene, S. (2017). Brain networks for confidence weighting and hierarchical inference during probabilistic learning. *Proceedings of the National Academy of Sciences*, *114*(19), E3859-E3868.
- Middlebrooks, P. G., & Sommer, M. A. (2012). Neuronal correlates of metacognition in primate frontal cortex. *Neuron*, *75*(3), 517-530.
- Milgram, S. (1974). The dilemma of obedience. *The Phi Delta Kappan*, *55*(9), 603-606.
- Morgan, T. J., Laland, K. N., & Harris, P. L. (2015). The development of adaptive conformity in young children: effects of uncertainty and consensus. *Developmental science*, *18*(4), 511-524.
- Morris, B. J., Croker, S., Masnick, A. M., & Zimmerman, C. (2012). The emergence of scientific reasoning. In *Current topics in children's learning and cognition*. IntechOpen.
- Morewedge, C. K., Yoon, H., Scopelliti, I., Symborski, C. W., Korris, J. H., & Kassam, K. S. (2015). Debiasing decisions: Improved decision making with a single training intervention. *Policy Insights from the Behavioral and Brain Sciences*, *2*(1), 129-140.
- Neill, W. T., Valdes, L. A., & Terry, K. M. (1995). Selective attention and the inhibitory control of cognition. In *Interference and inhibition in cognition* (pp. 207-261). Academic Press.
- Newman, I. R., & Thompson, V. A. (2017). Logical intuitions and other conundra for dual process theories. In *Dual Process Theory 2.0* (pp. 129-144). Routledge.
- Nisbett, R. E. (2015). *Mindware: Tools for smart thinking*. Farrar, Straus and Giroux.
- Nisbett, R. E., Fong, G. T., Lehman, D. R., & Cheng, P. W. (1987). Teaching reasoning. *Science*, *238*(4827), 625-631.
- Norman, E., Pfuhl, G., Sæle, R. G., Svartdal, F., Låg, T., & Dahl, T. I. (2019). Metacognition in psychology. *Review of General Psychology*, *23*(4), 403-424.

- Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, *101*(4), 608.
- Ott, T., Masset, P., & Kepecs, A. (2018, January). The neurobiology of confidence: from beliefs to neurons. In *Cold Spring Harbor symposia on quantitative biology* (Vol. 83, pp. 9-16). Cold Spring Harbor Laboratory Press.
- Palmer, E. C., David, A. S., & Fleming, S. M. (2014). Effects of age on metacognitive efficiency. *Consciousness and cognition*, *28*, 151-160.
- Pasquinelli, E. (2006). *An analysis of the notion of illusion and illusory phenomena. Illusions in haptic, dynamic, kinesthetic touch* (Doctoral dissertation).
- Pasquinelli, E. (2012). The awareness of illusions. In *Perceptual Illusions* (pp. 59-74). Palgrave Macmillan, London.
- Pennycook, G. (2017). A perspective on the theoretical foundation of dual-process models. *Dual process theory*, *2*, 34.
- Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015). Everyday consequences of analytic thinking. *Current Directions in Psychological Science*, *24*(6), 425-432.
- Pennycook, G., De Neys, W., Evans, J. S. B., Stanovich, K. E., & Thompson, V. A. (2018). The mythical dual-process typology. *Trends in Cognitive Sciences*, *22*(8), 667-668.
- Pennycook, G., Ross, R. M., Koehler, D. J., & Fugelsang, J. A. (2017). Dunning–Kruger effects in reasoning: Theoretical implications of the failure to recognize incompetence. *Psychonomic Bulletin & Review*, *24*(6), 1774-1784.
- Pohl, R. (2017). Cognitive illusions: Intriguing phenomena in judgment, thinking, and memory. *Abingdon, UK: Routledge*.
- Pouget, A., Beck, J. M., Ma, W. J., & Latham, P. E. (2013). Probabilistic brains: knowns and unknowns. *Nature neuroscience*, *16*(9), 1170.
- Pouget, A., Drugowitsch, J., & Kepecs, A. (2016). Confidence and certainty: distinct probabilistic quantities for different goals. *Nature neuroscience*, *19*(3), 366.
- Proust, J. (2013). *The philosophy of metacognition: Mental agency and self-awareness*. OUP Oxford.
- Proust, J. La métacognition. Les enjeux pédagogiques de la recherche. https://educamus.ac-versailles.fr/IMG/pdf/metacognition_gt5.pdf
- Rakoczy, H. (2017). In defense of a developmental dogma: Children acquire propositional attitude folk psychology around age 4. *Synthese*, *194*(3), 689-707.
- Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature neuroscience*, *2*(1), 79-87.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory*, *2*, 64-99.
- Richerson, P. J., & Boyd, R. (2008). *Not by genes alone: How culture transformed human evolution*. University of Chicago press.
- Rode, C., Cosmides, L., Hell, W., & Tooby, J. (1999). When and why do people avoid unknown probabilities in decisions under uncertainty? Testing some predictions from optimal foraging theory. *Cognition*, *72*(3), 269-304.
- Rouault, M., McWilliams, A., Allen, M. G., & Fleming, S. M. (2018). Human metacognition across domains: insights from individual differences and neuroimaging. *Personality neuroscience*, *1*.

- Rouault, M., Dayan, P., & Fleming, S. M. (2019). Forming global estimates of self-performance from local confidence. *Nature communications*, *10*(1), 1-11.
- Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive science*, *26*(5), 521-562.
- Shadlen, M. N., & Kiani, R. (2013). Decision making as a window on cognition. *Neuron*, *80*(3), 791-806.
- Sharot, T. (2011). The optimism bias. *Current biology*, *21*(23), R941-R945.
- Sharot, T., Korn, C. W., & Dolan, R. J. (2011). How unrealistic optimism is maintained in the face of reality. *Nature neuroscience*, *14*(11), 1475.
- Sharot, T., Kanai, R., Marston, D., Korn, C. W., Rees, G., & Dolan, R. J. (2012). Selectively altering belief formation in the human brain. *Proceedings of the National Academy of Sciences*, *109*(42), 17058-17062.
- Sharot, T., & Garrett, N. (2016). Forming beliefs: Why valence matters. *Trends in cognitive sciences*, *20*(1), 25-33.
- Schauble, L. (1990). Belief revision in children: The role of prior knowledge and strategies for generating evidence. *Journal of experimental child psychology*, *49*(1), 31-57.
- Schauble, L. (1996). The development of scientific reasoning in knowledge-rich contexts. *Developmental Psychology*, *32*(1), 102.
- Shea, N., Boldt, A., Bang, D., Yeung, N., Heyes, C., & Frith, C. D. (2014). Supra-personal cognitive control and metacognition. *Trends in cognitive sciences*, *18*(4), 186-193.
- Schulz, L. E., & Gopnik, A. (2004). Causal learning across domains. *Developmental psychology*, *40*(2), 162.
- Schulz, L. E., Gopnik, A., & Glymour, C. (2007). Preschool children learn about causal structure from conditional interventions. *Developmental science*, *10*(3), 322-332.
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593-1599.
- Serra, M. J., & Metcalfe, J. (2009). 15 Effective Implementation of Metacognition. *Handbook of metacognition in education*, 278.
- Scott, R. M., & Baillargeon, R. (2017). Early false-belief understanding. *Trends in Cognitive Sciences*, *21*(4), 237-249.
- Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., & Stine-Morrow, E. A. (2016). Do "brain-training" programs work?. *Psychological Science in the Public Interest*, *17*(3), 103-186.
- Sloman, S., & Fernbach, P. (2018). *The knowledge illusion: Why we never think alone*. Penguin.
- Smith, J. D., Shields, W. E., & Washburn, D. A. (2003). The comparative psychology of uncertainty monitoring and metacognition. *Behavioral and brain sciences*, *26*(3), 317-339.
- Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental science*, *10*(1), 89-96.
- Sperber, D. (1996). Explaining culture: A naturalistic approach. *Cambridge, MA: Cambridge*.
- Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. Mind & Language, 25(4), 359–393.

- Stanovich, K. E., & Stanovich, P. J. (2010). *A framework for critical thinking, rational thinking, and intelligence. Innovations in educational psychology: Perspectives on learning, teaching and human development*, 195-237.
- Stephens, D. W. (2008). Decision ecology: foraging and the ecology of animal decision making. *Cognitive, Affective, & Behavioral Neuroscience*, *8*(4), 475-484.
- Sutton, R. S., & Barto, A. G. (1998). *Introduction to reinforcement learning* (Vol. 135). Cambridge: MIT press.
- Tamis-LeMonda, C. S., Adolph, K. E., Lobo, S. A., Karasik, L. B., Ishak, S., & Dimitropoulou, K. A. (2008). When infants take mothers' advice: 18-month-olds integrate perceptual and social information to guide motor action. *Developmental psychology*, *44*(3), 734.
- Tipper, S. P. (2001). Does negative priming reflect inhibitory mechanisms? A review and integration of conflicting views. *The Quarterly Journal of Experimental Psychology Section A*, *54*(2), 321-343.
- Trippas, D., & Handley, S. J. (2018). The parallel processing model of belief bias: Review and extensions.
- Trouche, E., Sander, E., & Mercier, H. (2014). Arguments, more than confidence, explain the good performance of reasoning groups. *Journal of Experimental Psychology: General*, *143*(5), 1958.
- Trouche, E., Shao, J., & Mercier, H. (2019). Objective evaluation of demonstrative arguments. *Argumentation*, *33*(1), 23-43.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, *185*(4157), 1124-1131.
- van Bergen, Y., Coolen, I., & Laland, K. N. (2004). Nine-spined sticklebacks exploit the most reliable source when public and private information conflict. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *271*(1542), 957-962.
- Vo, V. A., Li, R., Kornell, N., Pouget, A., & Cantlon, J. F. (2014). Young children bet on their numerical skills: metacognition in the numerical domain. *Psychological science*, *25*(9), 1712-1721.
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta analysis of theory of mind development: The truth about false belief. *Child development*, *72*(3), 655-684
- Yaniv, I. (2004). Receiving other people's advice: Influence and benefit. *Organizational behavior and human decision processes*, *93*(1), 1-13.
- Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: confidence and error monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1310-1321.
- Zimmerman, C. (2000). The development of scientific reasoning skills. *Developmental review*, *20*(1), 99-149.
- Zimmerman, C., Raghavan, K., & Sartoris, M. (2003). The impact of the MARS curriculum on students' ability to coordinate theory and evidence. *International Journal of Science Education*, *25*(10), 1247-1271.

4. Éducabilité de L'EC

- Abrami, P. C., Bernard, R. M., Borokhovski, E., Waddington, D. I., Wade, C. A., & Persson, T. (2015). Strategies for teaching students to think critically: A meta-analysis. *Review of Educational Research*, *85*(2), 275-314.

- Adey, P., Shayer, M., & Yates, C. (1989). Thinking science: Student and teachers' materials for the CASE intervention. *London, Mccacmillan*.
- Anderson, J. R., Reder, L. M., & Simon, H. A. (1996). Situated learning and education. *Educational researcher*, *25*(4), 5-11.
- Bassok, M., & Olseth, K. L. (1995). Judging a book by its cover: Interpretative effects of content on problem-solving transfer. *Memory & Cognition*, *23*(3), 354-367.
- Biederman, I., & Shiffrar, M. M. (1987). Sexing day-old chicks: A case study and expert systems analysis of a difficult perceptual-learning task. *Journal of Experimental Psychology: Learning, memory, and cognition*, *13*(4), 640.
- Bjorklund, D. F., & Harnishfeger, K. K. (1990). Children's strategies: Their definition and origins. *Children's strategies: Contemporary views of cognitive development*, 309-323.
- Blagg, N. (1988). *SOMERSET Thinking Skills Course. Handbook + Modules 1-5: 1. Foundations for Problem Solving. 2. Analysing and Synthesising. 3. Comparative Thinking. 4. Positions in Time and Space. 5. Understanding Analogies*. Blackwell/Somerset County Council.
- Boku, M., Yama, H., & Mercier, H. (2018). Robust improvements on reasoning performance following discussion in Japan. *Japanese Psychological Research*, *60*(1), 47-53.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn* (Vol. 11). Washington, DC: National academy press.
- Brown, A. L., Palincsar, A. S., & Armbruster, B. B. (1984). Instructing comprehension-fostering activities in interactive learning situations. *Learning and comprehension of text*, 255-286.
- Brown, P. C., Roediger III, H. L., & McDaniel, M. A. (2014). *Make it stick*. Harvard University Press.
- Carraher, T. N. (1986). From drawings to buildings; working with mathematical scales. *International Journal of Behavioral Development*, *9*(4), 527-544.
- Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. *The psychology of learning and motivation*, *16*, 1-58.
- Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*, 55–81.
- Chi, M. T. H., & Feltovich, P. R. Glaser, and R. Rees. (1982). Expertise in problem solving. *Advances in the psychology of human intelligence*, *1*, 1-75.
- Chi, M. T., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive science*, *5*(2), 121-152.
- Chi, M. T., Glaser, R., & Farr, M. J. (1988). The nature of expertise. Hillsdale, NJ: L.
- Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. *The adapted mind: Evolutionary psychology and the generation of culture*, *163*, 163-228.
- de Groot, A. D. (1965). *Thought and choice in chess* (2nd ed. 1978). The Hague: Mouton.
- Dunker, K. (1945). On problem solving. Psych. *Monographs*, *58*, 270.
- Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. *Educational researcher*, *18*(3), 4-10.
- Ericsson, K. A., & Chase, W. G. (1982). *Exceptional memory* (No. TR-8-ONR). Carnegie-Mellon Univ Pittsburgh Pa Dept Of Psychology.

- Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological review*, *100*(3), 363.
- Facione, P. A. (1991). Using the California Critical Thinking Skills Test in Research, Evaluation, and Assessment.
- Facione, P. A. (2000). The disposition toward critical thinking: Its character, measurement, and relationship to critical thinking skill. *Informal logic*, *20*(1).
- Fayol, M., & Monteil, J. M. (1994). Note de synthèse. *Revue française de pédagogie*, *106*(1), 91-110.
- Feuerstein, R., Ya'acov Rand, Hoffman, M. B., & Miller, R. (1980). Instrumental enrichment: An intervention program for cognitive modifiability.
- Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive psychology*, *18*(3), 253-29
- Gamo, S., Sander, E., & Richard, J. F. (2010). Transfer of strategy use by semantic recoding in arithmetic problem solving. *Learning and Instruction*, *20*(5), 400-410.
- Gelder, T. V. (2005). Teaching critical thinking: Some lessons from cognitive science. *College teaching*, *53*(1), 41-48.
- Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology*, *95*(2), 393.
- Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive psychology*, *12*(3), 306-355.
- Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive psychology*, *15*(1), 1-38.
- Gobet, F., & Simon, H. A. (1996). Recall of random and distorted chess positions: Implications for the theory of expertise. *Memory & cognition*, *24*(4), 493-503.
- Halpern, D. F. (2013). *Thought and knowledge: An introduction to critical thinking*. Psychology Press.
- Hattie, J. A., & Donoghue, G. M. (2016). Learning strategies: A synthesis and conceptual model. *npj Science of Learning*, *1*, 16013.
- Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of educational research*, *66*(2), 99-136.
- Helfenstein, S., & Saariluoma, P. (2006). Mental contents in transfer. *Psychological Research*, *70*(4), 293-303.
- Helsdingen, A. S., Van den Bosch, K., Van Gog, T., & van Merriënboer, J. J. (2010). The effects of critical thinking instruction on training complex decision making. *Human factors*, *52*(4), 537-545.
- Hernstein, R. J., Nickerson, R. S., de Sanchez, M., & Swets, J. A. (1986). Teaching thinking skills. *American Psychologist*, *41*(11), 1279.
- Higgins, S., Hall, E., Baumfield, V., & Moseley, D. (2005). A meta-analysis of the impact of the implementation of thinking skills approaches on pupils.
- Higgins, S., Katsipataki, M., Villanueva-Aguilera, A. B., Coleman, R., Henderson, P., Major, L. E., ... & Mason, D. (2016). The Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit.
- Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & cognition*, *15*(4), 332-340.
- Kosonen, P., & Winne, P. H. (1995). Effects of teaching statistical laws on reasoning about everyday problems. *Journal of Educational Psychology*, *87*(1), 33.
- Kuhn, D. (2015). Thinking together and alone. *Educational researcher*, *44*(1), 46-53.

- Kuhn, D., Hemberger, L., & Khait, V. (2016). Tracing the development of argumentive writing in a discourse-rich context. *Written Communication*, *33*(1), 92-121.
- Kyndt, E., Raes, E., Lismont, B., Timmers, F., Cascallar, E., & Dochy, F. (2013). A meta-analysis of the effects of face-to-face cooperative learning. Do recent studies falsify or verify earlier findings? *Educational research review*, *10*, 133-149.
- Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, *208*(4450), 1335-1342.
- Lehman, D. R., Lempert, R. O., & Nisbett, R. E. (1988). The effects of graduate training on reasoning: Formal discipline and thinking about everyday-life events. *American Psychologist*, *43*(6), 431.
- Lipman, M. S., & Sharp, A. M. AM and Oscanyan, FS (1980). Philosophy in the classroom.
- Luchins, A. S., & Luchins, E. H. (1970). The effects of order of presentation of information and explanatory models. *The Journal of Social Psychology*, *80*(1), 63-70.
- Marin, L. M., & Halpern, D. F. (2011). Pedagogy for developing critical thinking in adolescents: Explicit instruction produces greatest gains. *Thinking Skills and Creativity*, *6*(1), 1-13.
- Marzano, R. J. (1998). A theory-based meta-analysis of research on instruction.
- Mercier, H. (2016). The argumentative theory: Predictions and empirical evidence. *Trends in Cognitive Sciences*, *20*(9), 689-700.
- Mercier, H., & Sperber, D. (2011). Argumentation: its adaptiveness and efficacy. *Behavioral and Brain Sciences*, *34*(2), 94.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, *63*(2), 81.
- Minervino, R. A., Olguín, V., & Trench, M. (2017). Promoting interdomain analogical transfer: When creating a problem helps to solve a problem. *Memory & cognition*, *45*(2), 221-232.
- Nisbett, C. J., Jepson, R. E., & Krantz, D. H. (1993). Inductive reasoning: Competence or skill? *Rules for reasoning*, 70.
- Nisbett, R. E. (2013). *Rules for reasoning*. Psychology Press.
- Nisbett, R. E., Fong, G. T., Lehman, D. R., & Cheng, P. W. (1987). Teaching reasoning. *Science*, *238*(4827), 625-631.
- Nisbett, R. E., Krantz, D. H., Jepson, C., & Kunda, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. *Psychological review*, *90*(4), 339.
- Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, memory, and cognition*, *14*(3), 510.
- Perkins, D. N., & Salomon, G. (1992). Transfer of learning. *International encyclopedia of education*, *2*, 6452-6457.
- Ramus, F. (2012). Au-delà de l'inné et de l'acquis. *Hors-série La recherche-jeux*, 18-20.
- Raynal, L., Clément, E., & Sander, E. (2017). Challenging the superficial similarities superiority account for analogical retrieval.
- Recht, D. R., & Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology*, *80*(1), 16.
- Resnick, L. B. (1987). *Education and learning to think*. National Academies.

- Sala, G., & Gobet, F. (2017). Does far transfer exist? Negative evidence from chess, music, and working memory training. *Current Directions in Psychological Science*, *26*(6), 515-520.
- Salomon, G., & Perkins, D. (1988). Teaching for transfer. *Educational Leadership*, *46*(1), 22-32.
- Salomon, G., & Perkins, D. N. (1987). Transfer of cognitive skills from programming: When and how? *Journal of educational computing research*, *3*(2), 149-169.
- Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanism of a neglected phenomenon. *Educational psychologist*, *24*(2), 113-142.
- Sills, J., Rowse, G., & Emerson, L. M. (2016). The role of collaboration in the cognitive development of young children: a systematic review. *Child: care, health and development*, *42*(3), 313-324.
- Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill* (No. 9). Harvard University Press.
- Slavin, R. E. (2014). Cooperative Learning and Academic Achievement: Why Does Groupwork Work? [Aprendizaje cooperativo y rendimiento académico: ¿por qué funciona el trabajo en grupo?] *Anales de psicología/annals of psychology*, *30*(3), 785-791.
- Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions: III. Functions involving attention, observation and discrimination. *Psychological Review*, *8*(6), 553.
- Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. *The adapted mind: Evolutionary psychology and the generation of culture*, *19*.
- Trouche, E., Sander, E., & Mercier, H. (2014). Arguments, more than confidence, explain the good performance of reasoning groups. *Journal of Experimental Psychology: General*, *143*(5), 1958.
- Van Gelder, T. (2001, December). How to improve critical thinking using educational technology. In *Meeting at the crossroads: Proceedings of the 18th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education* (pp. 539-548).
- Willingham, D. T. (2007). Critical thinking: Why it is so hard to teach? *American Federation of Teachers summer 2007, p. 8-19*.
- Willingham, D. (2019). How to teach critical thinking.
- Woodworth, R. S., & Thorndike, E. L. (1901). The influence of improvement in one mental function upon the efficiency of other functions.(I). *Psychological review*, *8*(3), 247.

Articles utilisés pour la méta-analyse (Abrami et al. 2105)

*****Abdulghani, B. A. (2003). *An inquiry into the effects of cooperative learning on critical thinking and achievement in the Arabic language by female high school students in the United Arab Emirates* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3114806)

*****Affeldt, J. E. O. (1994). *Occupational therapy clinical reasoning: supervision for the transition from student to clinician (critical thinking)* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9411515)

*****Al-Khayyatt, A. K. A. (1980). *An experimental study comparing the effects of the*

inquiry method and the traditional method for teaching social studies in two Kuwaiti secondary schools for boys (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8107532)

*****Allison, R. D. (1972). *An investigation into the attitudes toward science of college chemistry students as a function of laboratory experience.* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7300253)

*****Anderson, J. L. (1992). *A study of critical thinking as it relates to Myers-Briggs personality type, dominant function, and integrated process skills of ninth-grade physical science students.* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9400005)

*****Angeli, C., & Valanides, N. (2009). Instructional effects on critical thinking: Performance on ill-defined issues. *Learning and Instruction, 19*, 322-334. http://dx.doi.org/10.1016/j.learninstruc.2008.06.010

*****Annis, D. B., & Annis, L. (1979). An empirical study of the impact of philosophy on students' CT ability. *Teaching Philosophy, 3*, 145-152.

*****Apichatibutarapong, S., Logan, P., & Cobbin, D. (2004). Alternative strategies for teaching critical thinking. In J. Renner & J. Cross (Eds.), *New challenges in sustainability and growth in higher education: EDU-COM 2004 conference proceedings, 24-26 November 2004* (pp. 29-35). Joondalup, Western Australia: Edith Cowan University. Retrieved from

*****Arlo, R. (1968). *The relative effectiveness of inductive and expository teaching of principles of general semantics upon the critical reading ability of 9th-grade students.* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 6921178)

*****Armstrong, N. A. (1970). *The effect of two instructional inquiry strategies on critical thinking and achievement in eighth grade social studies* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7017235)

*****Arrufat, O. (1997). *The role of exploration and critical decision making and problem solving in making life choices* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9813418)

*****Asamani, J. B. B. (1998). *An experimental study of the relationship between cooperative learning and critical thinking* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9806590)

*****Aten, L. L. (2004). *Using guided learning journals to foster reflection and professional judgment in teachers* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3111581)

*****Athman, J. A. (2003). *The effects of environment-based education on students' critical thinking and achievement motivation* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3105582)

*****Ayoub, J. L. (1997). *The effect of classroom use of an electronic group support system on student critical thinking, performance and satisfaction* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9806760)

*****Bagby, L. A. (1991). *The study of popular culture to promote critical thinking in writing* (Master's thesis). Available from ProQuest Dissertations and Theses database. (UMI No. 1341542)

*****Bamvakais, K. M. (1981). *Promoting critical thinking: An instructional program for enhancing the remote transfer of critical thinking skills* (Doctoral dissertation).

Available from ProQuest Dissertations and Theses database. (UMI No. 8212317)
*****Bates, D. S. (1991). *Effect of a systematic inquiry intervention on the critical thinking skills of a sample of master's degree students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9206908)
*****Baumgartner, A. (1965). *An experimental study of speech fundamentals unit in public speaking upon organization skill and critical thinking ability* (Unpublished master's thesis). University of Nebraska, Lincoln.
*****Bean, A. T. (1969). *Some comparisons between conventional college teaching methods and a composite of procedures involving large lecture groups, seminars, and reduced class time* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 6919928)
*****Belinsky, S. B. (2000). *The effect of a peer to peer strategy within radiation therapy and nursing clinical settings on the development of critical thinking skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9961539)
*****Biddle, W. W. (1932). *Propaganda and education*. New York: AMS Press Inc.
*****Bird, J. J. (1984). *Effects of fifth graders' attitudes and critical thinking/reading skills resulting from a Junior Great Books program* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8411599)
*****Black, D. H. (1982). *Comparison of inquiry and non-inquiry methods in the development of critical thinking* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8208436)
*****Bonnette, R., McBride, R. E., & Tolson, H. (2001). The differential effect of indirect instruction in the teaching of sport skills on critical thinking and self-esteem of early adolescent boys placed at risk. *Sport, Education & Society, 6*, 183-198. http://dx.doi.org/10.1080/13573320120084263
*****Bosma, B. A. (1981). *An experimental study to determine the feasibility of using folk literature to teach select critical reading skills to sixth graders* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8126472)
*****Boyadjian-Samawi, Z. (2006). *The effect of concept mapping on critical thinking skills and dispositions of junior and senior baccalaureate nursing students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3207567)
*****Brownell, J. A. (1953). The influence of training in reading in the social studies on the ability to think critically. *California Journal of Educational Research, 4*, 28-31.
*****Bruning, R., Siwatu, K. O., Liu, X. Y., PytlikZillig, L. M., Horn, C., Sic, S., & Carlson, D. (2008). Introducing teaching cases with face-to-face and computer-mediated discussion: Two multi-classroom quasi-experiments. *Contemporary Educational Psychology, 33*(2), 299-326. http://dx.doi.org/10.1016/j.cedpsych.2006.11.002
*****Burkhart, L. M. (2006). *Thinking critically about critical thinking: Developing thinking skills among high school students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3214755)
*****Burns, E. (2009). *The use of science inquiry and its effect on critical thinking skills and dispositions in third grade students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3367974)
*****Butchart, S., Forster, D., Gold, I., Bigelow, J., Korb, K., Oppy, G., & Serrenti, A. (2009). Improving critical thinking using web based argument mapping exercises with

automated feedback. *Australasian Journal of Educational Technology, 25*, 268-291. Retrieved from http://www.ascilite.org.au/ajet/ajet25/butchart.pdf

*****Butchart, S., Handfield, T., & Restall, G. (2009). Using peer instruction to teach philosophy, logic, and critical thinking. *Teaching Philosophy, 32*(1), 1-40.

*****Butler, G. P. (1990). *A study of the effects of utilizing the writing process as an instructional strategy on student thinking skills in ninth-grade biology courses* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9030670)

*****Byrne, J. S. (1983). *The effect of critical thinking skills instruction on achievement and attitudes of elementary students differing in learning style preferences* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8417843)

*****Cabrera, G. A. (1982). *Evaluating the teaching of critical thinking in content areas* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8301629)

*****Carter, L. M. (2008). Critical thinking dispositions in online nursing education. *Journal of Distance Education, 22*(3), 89-114. Retrieved from http://www.jofde.ca/index.php/jde/article/view/454/772

*****Carwie, L. (2009). *The effect of computer-supported argument mapping on the critical thinking skills of undergraduate nursing students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3388766)

*****Chamberlain, M. A. (1993). *Philosophy for Children program and the development of critical thinking of gifted elementary students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9319950)

*****Champion, B. W. (1975). *An inductive approach to the discrimination of factual and inferential statements as an aspect of critical reading* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7528097)

*****Chen, F. C., & Lin, M. C. (2003). Effects of a nursing literature reading course on promoting critical thinking in two-year nursing program students. *Journal of Nursing Research, 11*, 137-147. http://dx.doi.org/10.1097/01.JNR.0000347629.80100.b2

*****Chennault, A. (1989). *Enhancing critical thinking skills in gifted elementary school students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8917052)

*****Clark, R. C. (1977). *The effects of graphic materials on critical thinking in social studies* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7726650)

*****Colbert, K. R. (1987). The effects of ceda and ndt debate training on critical thinking ability. *Journal of the American Forensic Association, 23*(4), 194-201.

*****Colledge, D. G. (1993). *An experimental study comparing embedding and immersion approaches to instructing critical thinking to 5th-grade students (fifth-grade)* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9334716)

*****Comerford, S. S., Busk, P. L., & Roberts, W. L. (2000, April). *Epistemological beliefs of community-college students*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED442483)

*****Conklin, K. E. (1986). *The impact of a critical thinking intervention on tenth grade students' expository writing skills* (Doctoral dissertation). Available from ProQuest

Dissertations and Theses database. (UMI No. 8629807)
*****Constantine, M. (1968). *An experimental study of the development of critical thinking skills of high school English teachers enrolled in a methods course* (Interim Report, Project No. HE-145). Illinois State-wide Curriculum Study Center in the Preparation of Secondary School English Teachers (ISCPET).
*****Cook, J. E. (1972). *A study in critical listening using eight to ten year olds in an analysis of commercial propaganda emanating from television* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7226840)
*****Crawford, J. A. (1976). *An investigation of the transfer of critical thinking skills from language arts to social studies* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7703254)
*****Criner, L. A. (1992). *Teaching thinking and reasoning: A study of critical thinking in adults* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9237328)
*****Crow, L. W., & Haws, S. G. (1985, April). *The effects of teaching logical reasoning upon students' critical thinking and science achievement*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, French Lick Springs, IN. (ERIC Document Reproduction Service No. ED255371)
*****Curtis, C. K. (1980). Developing critical thinking skills in nonacademic social studies classes. *Alberta Journal of Educational Research, 26*, 75-84.
*****Dale, P. M., Ballotti, D., Handa, S., & Zych, T. (1997). An approach to teaching problem solving in the classroom. *College Student Journal, 31*, 76-79.
*****Daodee, S. (2002). *The implementation of teaching strategies for critical thinking development in nursing education in Thailand: A concerns-based approach* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. C809151)
*****Desnouee, J. R. (2000). *The influence of the artforming technique on the self-esteem and reading behaviors of at-risk first-year college students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9980926)
*****Dewey, J., & Bento, J. (2009). Activating children's thinking skills (ACTS): The effects of an infusion approach to teaching thinking in primary schools. *British Journal of Educational Psychology, 79*, 329-351. http://dx.doi.org/10.1348/000709908X344754
*****Diket, M. R. M. (1991). *Art criticism: Relationships to critical thinking, appreciation, and creativity among the gifted* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9206946)
*****Dixon, F. A. (1996). *The use of the Hegelian dialectic as a heuristic in literature classes to improve critical thinking skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9725536)
*****Dolan, S. L. (1976). *The effect of value-conflict discussion on the critical thinking and reading of adults* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7626218)
*****Donawa, A. (2009). *Critical thinking instruction and minority engineering students at a public urban higher education institution* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3396399)
*****Douglas, J. (1951). An experimental study of training in problem solving methods. *Speech Monographs, 19*, 115-116.
*****Downing, K., Kwong, T., Chan, S.-W., Lam, T.-F., & Downing, W.-K. (2009).

Problem-based learning and the development of metacognition. *Higher Education, 57*, 609-621. http://dx.doi.org/10.1007/s10734-008-9165-x

*****Drews, E. M. (1965). *Being and becoming--a cosmic approach to counseling and curriculum, part 2--the creative intellectual style in gifted adolescents*. East Lansing: Michigan State University.

*****Duatepe-Paksu, A., & Ubuz, B. (2009). Effects of drama-based geometry instruction on student achievement, attitudes, and thinking levels. *Journal of Educational Research, 102*, 272-286. http://dx.doi.org/10.3200/JOER.102.4.272-286

*****Duckworth, J. B. (1968). *The effect of instruction in general semantics on the critical thinking of tenth and eleventh grade students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 6906092)

*****Dukess, S. L. (1986). *Teaching literary skills to eighth graders through critical thinking: A theoretical model and a comparison of two approaches* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8522954)

*****Duphorne, P. L., & Gunawardena, C. N. (2005). The effect of three computer conferencing designs on critical thinking skills of nursing students. *American Journal of Distance Education, 19*, 37-50. http://dx.doi.org/10.1207/s15389286ajde1901_3

*****Eason, L. E. (1986). *The relationship of critical thinking skills and psychological type in community college students' responses to science instruction* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8704157)

*****Eimers, S. P. (1987). *Effects of reflective thinking instruction taught through the philosophy for children program on the reasoning ability and attitude toward teaching thinking of preservice elementary teachers* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8809971)

*****Elliott, B. M. (1999). *The influence of an interdisciplinary course on critical thinking skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9945800)

*****Endler, L. C., & Bond, T. G. (2008). Changing science outcomes: Cognitive acceleration in a US setting. *Research in Science Education, 38*, 149-166. http://dx.doi.org/10.1007/s11165-007-9042-0

*****Ennis, T. R. (2007). *Assessing and enhancing critical thinking skills: Enhanced Peer Review (EPR) with the CAT instrument* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3263266)

*****Ersözlü, Z. N., & Arslan, M. (2009). The effect of developing reflective thinking on metacognitional awareness at primary education level in Turkey. *Reflective Practice, 10*, 683-695. http://dx.doi.org/10.1080/14623940903290752

*****Estrada, J. M. (1992). *Effects of programming in basic and logo on problem-solving skills* (Master's thesis). Available from ProQuest Dissertations and Theses database. (UMI No. 1349192)

*****Fang, R. J., Lin, C. C., Yang, H. J., Lee, C. J., Tsai, H. L., & Tsai, T. S. (2008). A study to increase the critical thinking and problem solving abilities by web-based learning. In Q. Li, S. Y. Chen, A. Xu & M. Li (Eds.), *Musp '08: Multimedia Systems and Signal Processing* (pp. 75-80).

*****Faryniarz, J. V. (1990). *The effectiveness of microcomputer simulators to stimulate environmental problem-solving with community college students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No.

9002473)
*****Faryniarz, J. V., & Lockwood, L. G. (1992). Effectiveness of microcomputer simulations in stimulating environmental problem solving by community college students. *Journal of Research in Science Teaching, 29*, 453-470. http://dx.doi.org/10.1002/tea.3660290503
*****Fennelly, P. (1989). *The effects of one program in critical thinking instruction on the reading comprehension ability of underprepared college students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8813592)
*****Ferrell, S. D. (1992). *Critical thinking as a function of teacher questions* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9206949)
*****Feuerstein, M. (2002). *Media literacy in support of critical thinking* (Unpublished doctoral dissertation). Liverpool, England.
*****Fiasca, M. A. (1966). *Feasibility of integration of selected aspects of (CBA) Chemistry, (CHEMS) Chemistry and (PSSC) Physics into a two year Physical Science sequence*. Ann Arbor, MI. University Microfilms.
*****Fielding, R. (1985). A comparative study of instructional modes for developing visual art criticism skills. *Journal of the Institute of Art Education, 9*(3), 19-28.
*****Fischer, C. W. (1990). *Effects of a developmental drama-inquiry process on creative and critical thinking skills in early adolescent students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9010183)
*****Fisher, D. E. L. (1968). *An experimental study of the teaching of certain critical thinking skills to ninth grade pupils* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 6814977)
*****Forbes, C. A. (1997). *Analyzing the growth of the critical thinking skills of college calculus students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9737709)
*****Frank, A. (1964). *An experiment in teaching critical thinking in a high school speech fundamentals course* (Unpublished doctoral dissertation). University of Wisconsin, Madison.
*****Frank, A. D. (1969). Teaching high school speech to improve critical thinking ability. *The Speech Teacher, 18*, 297-302.
*****Freeman, B. (1995). *A study of the efficacy of a developmental instruction design for non-traditional adult learners (nontraditional students)* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9432514)
*****Friend, H. A. (1967). *A comparison of the relative effectiveness of two methods of teaching the course Time, Space, and Matter to selected eighth grade pupils* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 6911745)
*****Gadzella, B. M., Hartsoe, K., & Harper, J. (1989). Critical thinking and mental ability groups. *Psychological Reports, 65*(3, Pt 1), 1019-1026.
*****Galinski, M. E. (1989). *An investigation of the relationship of microcomputer problem-solving software and critical thinking abilities of junior-high students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8818884)
*****Gent, P. J. (1997). *Using case studies in a course on disabilities to improve the attitudes toward inclusion and the critical thinking skills of preservice educators*

(Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9716983)

*****Germann, P. J. (1989). Directed-inquiry approach to learning science process skills: treatment effects and aptitude-treatment interactions. *Journal of Research in Science Teaching, 26*, 237-250. http://dx.doi.org/10.1002/tea.3660260305

*****Ghee, H. J. (1976). *A study of the effects of high level cognitive questions on the levels of response and critical thinking abilities in students of two social problems classes* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7601090)

*****Gibbs, L. E., Browne, M. N., & Keeley, S. M. (1988). *Stimulating critical thinking through faculty development: design, evaluation, and problems*. Washington, DC: American Association of State Colleges and Universities, Wisconsin University.

*****Gibson, H. W. (1986). *Critical thinking: A communication model* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8601471)

*****Giles, K. S. (1990). *Improving students' transactional writing through cognitive strategy intervention* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9008808)

*****Gillespie, M. A. (2005). *Critical thinking about values: The effects of an instructional program, reasons for attending college, and general life goals on the application of critical thinking to values expressed in an essay prompt* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3193386)

*****Glaser, E. M. (1941). *An experiment in the development of critical thinking*. New York: AMS Press.

*****Griffin, C. M. (1979). *The effect of the use of selected newspaper articles in the teaching of certain critical reading/thinking skills* (Unpublished M.Ed. Thesis). Rutgers University.

*****Gudaitis, D. J. (1971). *The effects of two seventh grade science programs, "interaction of man and the biosphere" and "science is explaining", on student attitude, science processes, and critical thinking* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7124823)

*****Guice, E. D. (1992). *The effect of instruction in concept analysis on critical thinking skills and moral reasoning decisions of senior baccalaureate nursing students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9225819)

*****Hagan, M. W. (2005). *Critical thinking skills study: Vocational nursing* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3174552)

*****Hagelskamp, J. L. (2001). *The effect of teaching critical thinking by infusion, with focus on transfer of skills, on the moral judgment and critical thinking of secondary school students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9988873)

*****Hamilton, L. J. (1994). *Relationship of critical thinking and learning styles to nursing diagnosis* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9411644)

*****Hancock, B. W., Jr. (1982). *The effect of guided design on the critical thinking ability of college level administrative science students* (Doctoral dissertation). Available from

ProQuest Dissertations and Theses database. (UMI No. 8206458)
*****Hancock, B. W., Coscarelli, W. C., & White, G. P. (1983). Critical thinking and content acquisition using a modified guided design process for large course sections. *Educational and Psychological Research, 3*(3), 139-149.
*****Hanninen, G. E. (1989). *The effects of the Hilda Taba teaching strategies on critical and creative thinking* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8918535)
*****Harbeck, C. A. (1973). *Experimental study of the effect of two proof formats in high school geometry on critical thinking and selected student attitudes* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. AAT 7302013)
*****Harrell, M. (2011). Argument diagramming and critical thinking in introductory philosophy. *Higher Education Research & Development, 30*(3), 371-385. Retrieved from http://www.hss.cmu.edu/philosophy/harrell/HarrellHERDSubNotReview.pdf
*****Harty, H., Woods, C. J., Johnson, D. A., & Pifer, S. (1986). *Academic challenge program: evaluation report*. Indianapolis, IN.: Indiana State Dept. of Education.
*****Hash, R. J. (1974). *The effects of a strategy of structured overviews, levels guides and vocabulary exercises on student achievement, reading comprehension, critical thinking and attitudes of junior high school classes in social studies* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7429239)
*****Henderson, J. S. (2002). *Critical to your health: using the theory of planned behavior to teach critical thinking skills and their effect on health behaviors* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3053289)
*****Hendrix, N. H. (1995). *Improving critical thinking and reading achievement in community college students.* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9500238)
*****Heraty, N., & Morley, M. J. (2000). The application of the structure of intellect programme: A manufacturing facility experiment. *Journal of Managerial Psychology, 15*(7), 691-715. http://dx.doi.org/10.1108/02683940010378063
*****Hernstein, R., Nickerson, R., Sanchez, M., & Swets, J. (1986). Teaching thinking skills. *American Psychologist, 41*, 1279-1289. http://dx.doi.org/10.1037/0003-066X.41.11.1279
*****Hesterberg, L. J. (2005). *Evaluation of a problem-based learning practice course: Do self-efficacy, critical thinking, and assessment skills improve?* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3162941)
*****Hill, N. J. (1997). *Impact of a critical thinking continuing education program on the critical thinking skills of practicing nurses* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9737644)
*****Hinton, R. S. (1980). *A study of the effects of utilizing local history on retention of historical information, critical thinking skills, and attitudes in teaching United States history.* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8010670)
*****Horn, C., & Bruning, R. (2004). The role of online discussion in promoting critical thinking using: Case studies with pre-service teachers. In L. Cantoni & C. McLoughlin (Eds.), *Proceedings of World Conference on Educational Multimedia,*

Hypermedia and Telecommunications 2004 (pp. 3088-3094). Chesapeake, VA: AACE. Retrieved from http://www.editlib.org/p/12459

*****Housen, A. C. (2001-2002). Aesthetic thought, critical thinking and transfer. *Arts and Learning Research Journal*, *18*(1), 99-132.

*****Houtveen, A. A. M., & Grift, W. J. C. M. v. d. (2007). Effects of metacognitive strategy instruction and instruction time on reading comprehension. *School Effectiveness and School Improvement*, *18*, 173-190. http://dx.doi.org/10.1080/09243450601058717

*****Howard, V. M. (2007). *A comparison of educational strategies for the acquisition of medical-surgical nursing knowledge and critical thinking skills: Human patient simulator vs. the interactive case study approach* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3270096)

*****Hsu, L. L. (2004). Developing concept maps from problem-based learning scenario discussions. *Journal of Advanced Nursing*, *48*, 510-518. http://dx.doi.org/10.1111/j.1365-2648.2004.03233.x

*****Hunt, C. A. (2003). *Promoting critical thinking: an analysis of instructional techniques with undergraduates in a university setting* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3076096)

*****Hur, S.-J. (2001). *Effects of anchored instruction on the critical-thinking skills of students with and without mild disabilities* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3037504)

*****Inlow, F. H., & Chovan, W. (1993). Another search for the effects of teaching thinking and problem solving skills on college students' performance. *Journal of Instructional Psychology*, *20*, 215-223.

*****Janiszewski Goodin, H. I. (2005). *The use of deliberative discussion as a teaching strategy to enhance the critical thinking abilities of freshman nursing students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3176411)

*****Johnson, A. P. (1997). *The effect of general thinking skills instruction on the thinking and writing performance of fifth and sixth grade students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9719668)

*****Jones, C. (1970). *The effects of participation in discussion of reasoning in arguments on critical thinking ability* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7013719)

*****Jones, H. E. (1993). *The effect of direct instruction of thinking skills in elementary social studies on the development of thinking skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9322839)

*****Jones, J. T. (1975). *An experimental study of four interdisciplinary approaches to promoting critical thinking skills and personal development in the college classroom* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7503500)

*****Joung, S. (2004). *The effects of high-structure cooperative versus low-structure collaborative design on online debate in terms of decision making, critical thinking, and interaction pattern* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3109292)

*****Joyce, J. (1971). *A comparison of two methods of teaching life career planning to junior high school students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7204087)

*****Kaberman, Z., & Dori, Y. J. (2009). Question posing, inquiry, and modeling skills of chemistry students in the case-based computerized laboratory environment. *International Journal of Science and Mathematics Education, 7*, 597-625. http://dx.doi.org/10.1007/s10763-007-9118-3

*****Kaberman, Z., & Dori, Y. J. (2009). Metacognition in chemical education: question posing in the case-based computerized learning environment. *Instructional Science, 37*, 403-436. http://dx.doi.org/10.1007/s11251-008-9054-9

*****Kaplan, I. S. (1997). *An assessment of the infusion of critical thinking skills into content instruction* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9737548)

*****Keller, R. (1994). *Effects of an instructional program on critical thinking and clinical decision-making skills of associate degree nursing students (nursing education)* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9404905)

*****Kemp, S. G., & Sadoski, M. (1991). The effects of instruction in forming generalizations on high school students' critical thinking in world history. *Reading Research and Instruction, 31*(1), 33-42.

*****Khoiny, F. E. (1996). *The effectiveness of problem-based learning in nurse practitioner education* (Doctoral dissertations). Available from ProQuest Dissertations and Theses database. (UMI No. 9614036)

*****Kirschbaum, H. (1979). *The effects of presenting general semantics lessons in five media upon the development of critical thinking in junior high school students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7925272)

*****Klauer, K. J. (1996). Teaching inductive reasoning: some theory and three experimental studies. *Learning and Instruction, 6*, 37-57. http://dx.doi.org/10.1016/S0959-4752(96)80003-X

*****Knight, D. H. (1981). *The effects of transfer materials on the critical thinking abilities of second-year algebra students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8107220)

*****Knoff, E., Rugh, J. D., & Littlefield, J. H. (1990). Comparing learning logs and note summaries in a dental research methods course. *Journal of Dental Education, 54*, 724-729.

*****Knotts, J. D. (1998). *Differential effectiveness of two instructional methods for developing higher-order question writing skills of preservice special educators* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9808829)

*****Koehler, C. F. (1994). *The effectiveness of deconstructive teaching strategies on students' critical thinking skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9409422)

*****Koray, O., & Koksal, M. S. (2009). The effect of creative and critical thinking based laboratory applications on creative and logical thinking abilities of prospective teachers. *Asia-Pacific Forum on Science Learning and Teaching, 10*(1). Retrieved from http://www.ied.edu.hk/apfslt/v10_issue1/koksal/

*****Korkmaz, O., & Karakus, U. (2009). The impact of blended learning model on student attitudes towards geography course and their critical thinking dispositions and levels. *Turkish Online Journal of Educational Technology, 8*(4), 51-63.

*****Kowalski, S. D., & Louis, M. A. (2000). Using computer technology to teach critical

thinking. *Nurse Educator, 25*, 210-212.
*****Kramarski, B., & Zoldan, S. (2008). Using errors as springboards for enhancing mathematical reasoning with three metacognitive approaches. *Journal of Educational Research, 102*, 137-151.
*****Krank, H. M. (1994). *Do cognitive styles and learning environment predict critical thinking performance for pre-service teachers?* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9405077)
*****Kumta, S., Tsang, P., Hung, L., & Cheng, J. (2003). Fostering critical thinking skills through a web-based tutorial programme for final year medical students – a randomized controlled study. *Journal of Educational Multimedia and Hypermedia, 12*, 267-273.
*****Langer, P., & Chiszar, D. (1993). Assessment of critical thinking courses. *Perceptual and Motor Skills, 77*, 970.
*****Lee, C., Tsai, H., & Lin, C. (2003). A Web-based Instructional Design for Teaching Critical Thinking Ability---Using Current Issues As An Example. In C. Crawford et al. (Ed.), *Proceedings of the Society for Information Technology and Teacher Education International Conference 2003* (pp. 3281-3284). Chesapeake, VA: AACE.
*****Lee, K. (2005). Enhancing critical thinking in online learning. *Academic Exchange Quarterly, 9*(4), 43-48.
*****Lee, K. (2007). Online collaborative case study learning. *Journal of College Reading and Learning, 37*(2), 82-100.
*****Lee, M. O. C., & Thompson, A. (1997). Guided instruction in logo programming and the development of cognitive monitoring strategies among college students. *Journal of Educational Computing Research, 16*, 125-144. http://dx.doi.org/10.2190/PW3F-HLFD-1NNJ-H77Q
*****Leonard, M. (2007). *Teaching critical thinking in a community-college art-history learning environment: A comparison of implicit instruction versus a combination of implicit and explicit instruction* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3280031)
*****Lewis, H. (1962). *An experiment in developing critical thinking through the teaching of plane demonstrative geometry* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 0001926)
*****Lin, C.-C., Lee, C.-J., & Tsai, H.-L. (2003). A web-based instructional design for teaching critical thinking ability using current issues as an example. In C. Crawford et al. (Ed.), *Proceedings of the Society for Information Technology and Teacher Education International Conference* (pp. 3281-3284). Chesapeake, VA: AACE. Retrieved from http://www.editlib.org/p/18693
*****Lindsay, C. W. (1974). *A study of the effects of three methods of teaching high school chemistry upon achievement in chemistry, critical thinking abilities, and scientific interest* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7411454)
*****Little, C. A., Feng, A. X., VanTassel, B. J., Rogers, K. B., & Avery, L. D. (2007). A study of curriculum effectiveness in social studies. *Gifted Child Quarterly, 51*(3), 272-284.
*****Liu, C.-C. (2006). *Connecting the dots between critical thinking skills and dispositions for middle school students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3243081)
*****Liu, M., & Pedersen, S. (1998, June). *The effect of being hypermedia designers on*

elementary school students' motivation and learning of design knowledge. Paper presented at the ED-MEDIA/ED-TELECOM 98 World Conference on Educational Multimedia and Hypermedia & World Conference on Educational Telecommunications, Freiburg, Germany. (ERIC Document Reproduction Service No. ED428689)

*****Livingston, H. (1964). *The effect of instruction in general semantics on the critical reading ability of 10th grade students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 6506581)

*****Lohse, B., Nitzke, S., & Ney, D. M. (2003). Introducing a problem-based unit into a lifespan nutrition class using a randomized design produces equivocal outcomes. *Journal of the American Dietetic Association, 103*, 1020-1025. http://dx.doi.org/10.1016/S0002-8223(03)00478-4

*****Loiacono, C. M. (2000). *Promoting freshman college student development using cognitive developmental theory presented in a deliberate psychological education-based freshman orientation program* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9974946)

*****Loveless, E. M. (1970). *Developing critical thinking skills in students in eighth grade English classes* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7013720)

*****Luckett, R. A. (1991). *The effect of the use of algorithmization on critical thinking skills of upper-division college students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9127000)

*****Lumpkin, C. R. (1991). *Effects of teaching critical thinking skills on the critical thinking ability, achievement, and retention of social studies content by fifth and sixth-graders (fifth- graders)* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9109097)

*****Lyons, E. B. (2006). *Examining the effects of problem-based learning on the critical thinking skills of associate degree nursing students in a southeastern community college* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3227670)

*****Lyons, E. M. (2008). Examining the effects of problem-based learning and NCLEX-RN scores on the critical thinking skills of associate degree nursing students in a Southeastern Community College. *International Journal of Nursing Education Scholarship, 5*, http://dx.doi.org/10.2202/1548-923X.1524

*****Macklem, R. E. (1974). *The effects of a specific teacher inservice program in BSCS biology upon teacher and student verbal behaviors and selected student learning outcomes.* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7413000)

*****MacPhail Wilcox, B., Dreyden, J., & Eason, E. (1990). An investigation of paideia program effects on students' critical thinking skills. *Educational Considerations, 17*(2), 61-67.

*****Maina, M. P. (1997). *Effects of divergent production style h initiative game activities on fifth-grade students use of critical thinking skills in physical education* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9725010)

*****Mandell, A. (1967). *A study of the effects of laboratory experiences upon critical thinking ability in a college biology course* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 6701026)

*****Manning, W. J. H. (1999). *The relationship between critical thinking and attitudes toward reading of the community college student enrolled in a critical reading course at Roane State Community College* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9903929)

*****Martin, C. L. (2003). *The effect of training on students' critical thinking skills in computer conferences over time* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3061096)

*****Massey, S. L. (2004). *The effects of cooperative learning versus traditional classroom instruction on cognitive achievement, critical thinking, and attitudes toward learning in teams in a physician assistant program* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3098166)

*****Mazer, J. P., Hunt, S. K., & Kuznekoff, J. H. (2007). Revising general education: Assessing a critical thinking instructional model in the basic communication course. *Journal of General Education 56*, 173-199

*****Mcaulay, D. L. (1988). *Type ii enrichment by gifted/talented teachers: the effects of inservicing critical thinking on third-grade teachers and students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8728889)

*****McBride, R., & Bonnette, R. (1995). Teacher and at-risk students' cognitions during open-ended activities: structuring the learning environment for critical thinking. *Teaching and Teacher Education, 11*, 373-388. http://dx.doi.org/10.1016/0742-051X(94)00040-D

*****McCarthy-Tucker, S. N. (1995). *Teaching reality-based formal logic to adolescents to improve critical thinking skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9530227)

*****McCarthy-Tucker, S. N. (1998). Teaching logic to adolescents to improve thinking skills. *Korean Journal of Thinking and Problem Solving, 8*(1), 45-66.

*****McCormick, G. J. (1989). *The effects of using print media in teaching critical thinking skills to adult students.* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8901471)

*****McCrink, C. L. S. (1999). *The role of innovative teaching methodology and learning styles on critical thinking* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9905019)

*****Mcgregor, G. D. J. (2002). *Creative thinking instruction for a college study skills program: A case study* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3027933)

*****McKee, M. L. (2006). *Overcoming obstacles in teaching college students how to think critically: The effects of an instructional intervention* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3222188)

*****McKeon, L. M., Norris, T., Cardell, B., & Britt, T. (2009). Developing patient-centered care competencies among prelicensure nursing students using simulation. *Journal of Nursing Education, 48*(12), 711-715. http://dx.doi.org/10.3928/01484834-20091113-06

*****Mcmurray, M. A. (1989). *Biology analytical problem solving interventions and field independence as predictors of critical thinking changes* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8902756)

*****Meiss, G. T., & Bates, G. W. (1984, May). *Cognitive and attitudinal effects of reasoning message strategies*. Paper presented at the 34th annual meeting of the

International Communication Association, San Francisco, CA. (ERIC Document Reproduction Service No. ED246519)

*****Miri, B., Ben-Chaim, D., & Zoller, U. (2007). Purposely teaching for the promotion of higher-order thinking skills: A case of critical thinking. *Research in Science Education, 37*, 353-369. http://dx.doi.org/10.1007/s11165-006-9029-2

*****Montean, J. J. (1960). *An experimental study of the use of discussion groups in general chemistry and general science as a means of increasing group growth in critical thinking* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 5906311)

*****Montuori, L. A., & Kimmel, E. B. (1994, August). *Teaching conceptual complexity to adults using an in-basket instructional design.* Paper presented at the annual meeting of the American Psychological Association, Los Angeles, CA. (ERIC Document Reproduction Service No. ED388768)

*****Moreyra, A. A. (1992). *The role of thinking frames in developing teachers' critical thinking skills and dispositions* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9136492)

*****Neimark, E. D. (1984, August). *A cognitive style change approach to the modification of thinking in college students*. Paper presented at the Conference on Thinking, Cambridge, MA. (ERIC Document Reproduction Service No. ED261301)

*****Nielsen, M. E. (1985). *Evaluation of a rural gifted program: Assessment of attitudes, self-concepts, and critical thinking skills of high-ability students in grades 3 through 12* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8500417)

*****Nieto, A. M., & Saiz, C. (2008). Evaluation of Halpem's "structural component" for improving critical thinking. *The Spanish journal of psychology, 11*, 266-274. Retrieved from http://www.ucm.es/info/Psi/docs/journal/v11_n1_2008/art266.pdf

*****Noland, R. L. (2002). *The effectiveness of the ropes challenge course on the enhancement of critical thinking skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3055822)

*****Overton, J. C. (1994). *An investigation of the effects of thinking skills instruction on academic achievement and the development of critical and creative thinking skills of second-, fourth-, and sixth-grade students (second-grade, fourth- grade)* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9417153)

*****Pachtman, A. B. (1977). *The effects of a reading and language arts program on the critical thinking and critical reading of the first-year law students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7724569)

*****Parker, M. J. (1999). *The effects of a shared, intranet science learning environment on the academic behaviors of problem-solving and metacognitive reflection* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9912894)

*****Parkinson, M. G., & Ekachai, D. (2002). The Socratic method in the introductory PR course: An alternative pedagogy. *Public Relations Review, 28*, 167-174. http://dx.doi.org/10.1016/S0363-8111(02)00123-6

*****Pellegrino, A. M. (2007). *The manifestation of critical thinking and metacognition in secondary American history students through the implementation of lesson plans and activities consistent with historical thinking skills* (Doctoral dissertation). Available

from ProQuest Dissertations and Theses database. (UMI No. 3282653)
*****Penningroth, S. L., Despain, L. H., & Gray, M. J. (2007). Topical articles: A course designed to improve psychological critical thinking. *Teaching of Psychology, 34*, 153-157. http://dx.doi.org/10.1080/00986280701498509
*****Perkins, H. W. (1985). *The effects of microcomputer use on the critical thinking skills of middle school students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8506714)
*****Pfeffer, A. C. (1977). *The effect of an independent study process on the critical thinking ability of seventh grade middle school students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7713522)
*****Phelps, P. H. (1988). *The effects of participation in reflective teaching on preservice teachers' critical thinking* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8725303)
*****Pickens, C. L. (2007). *Concept mapping: Methods to improve critical thinking* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3279742)
*****Pierce, W., Lemke, E., & Smith, R. (1988, April). *Effects of a secondary curriculum for moral development and critical thinking.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED296940)
*****Pierleoni, R. G. (1968). *The affective measurement of a programmed unit on American political attitudes and positions* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 6815879)
*****Pisano, R. C. (1980). *The effectiveness of an intervention study in critical thinking skills designed to improve written composition in eleventh and twelfth graders* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8014250)
*****Poel, R. H. (1971). *Critical thinking as related to PSSC and non-PSSC physics programs* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7103944)
*****Quitadamo, I. J. (2003). *Critical thinking in higher education: The influence of teaching styles and peer collaboration on science and math learning* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3069650)
*****Quitadamo, I. J., Faiola, C. L., Johnson, J. E., & Kurtz, M. J. (2008). Community-based inquiry improves critical thinking in general education biology. *CBE - Life Sciences Education, 7*, 327-337. http://dx.doi.org/10.1187/cbe.07-11-0097
*****Quitadamo, I. J., & Kurtz, M. J. (2007). Learning to improve: Using writing to increase critical thinking performance in general education biology. *CBE Life Sciences Education, 6*, 140-154. http://dx.doi.org/10.1187/cbe.06-11-0203
*****Ray, C. L. (1979). *A comparative laboratory study of the effects of lower level and higher level questions on students' abstract reasoning and critical thinking in two non-directive high school chemistry classrooms* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7926803)
*****Reed, J. H. (1999). *Effect of a model for critical thinking on student achievement in primary source document analysis and interpretation, argumentative reasoning, critical thinking dispositions, and history content in a community college history course* (Doctoral dissertation). Available from ProQuest Dissertations and Theses

database. (UMI No. 9911510)
*****Reed, J. H., & Kromrey, J. D. (2001). Teaching critical thinking in a community college history course: Empirical evidence from infusing Paul's model. *College Student Journal, 35*, 201-215.
*****Reid, E. C. (1982). *Training higher-level cognitive skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8215823)
*****Renaud, R. D. (2002). *The effect of higher-order questions on critical thinking skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. NQ67183)
*****Resnick, L. B., Bill, V., & Lesgold, S. (1992). Developing thinking abilities in arithmetic class. In A. Demetriou, M. Shayer & A. Efklides (Eds.), *Neo-Piagetian Theories of Cognitive Development* (pp. 210-230). London: Routledge. Retrieved from
*****Rickert, R. K. (1967). Developing critical thinking. *Science Education, 51*, 24-27. http://dx.doi.org/10.1002/sce.3730510106
*****Riley, V. M. (1992). *Teachers' questioning for improvement of critical thinking skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9135488)
*****Rivers, R. H., & Vockell, E. (1987). Computer simulations to stimulate scientific problem solving. *Journal of Research in Science Teaching, 24*, 403-415. http://dx.doi.org/10.1002/tea.3660240504
*****Robinson, T. Y. (2005). *A study of the effectiveness of environmental education curricula in promoting middle school students' critical thinking skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3195340)
*****Rogers, G., & McDonald, L. (1992). Thinking critically - an approach to field instructor training. *Journal of Social Work Education, 28*, 166-177.
*****Rose, M. M. (1998). *Critical thinking skills instruction for postsecondary students with and without learning disabilities: The effectiveness of icons as part of a literature curriculum* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9806188)
*****Ross, G. A., & Semb, G. (1981). Philosophy can teach critical thinking skills. *Teaching Philosophy, 4*, 111-122.
*****Rothert, L. F. (1969). *An analysis of changes in critical thinking ability, open-mindedness, and farm policy opinions of participants in the Kellogg farmers study program* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 6920923)
*****Rothstein, A. (1962). *An experiment in developing critical thinking through the teaching of American history* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 6003758)
*****Ruff, L. G. (2005). *The development of critical thinking skills and dispositions in first-year college students: Infusing critical thinking instruction into a first-year transitions course* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3201999)
*****Ryan, L. G. (1996). *Critical thinking in social work practice: a quasi- experimental investigation* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9622836)
*****Sanders, J. A., Wiseman, R. L., & Gass, R. H. (1994). Does teaching argumentation facilitate critical thinking? *Communication Reports, 7*, 27-35.

http://dx.doi.org/10.1080/08934219409367580
*Sasscer, J. C. (1973). *The development, implementation, and evaluation of a modularized, student-centered general biology curriculum at the college level* (Unpublished doctoral dissertation). The Catholic University of America.
*Schroeder, J. M. (2007). *A study of improving critical thinking skills with multiple choice tests and first semester associate degree nursing students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3278154)
*Schuler, G. (1974, April). *The effectiveness of the productive thinking program.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, Illinois. (ERIC Document Reproduction Service No. ED103479)
*Schulhauser, C. E. (1991). *The effects of literary discussion groups on students' critical thinking ability and attitude toward reading* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9131095)
*Schulze, F. R. (1998). *The effects of instructional strategies on levels of abstract thought* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9815686)
*Sendag, S., & Odabasi, H. F. (2009). Effects of an online problem based learning course on content knowledge acquisition and critical thinking skills. *Computers & Education, 53*, 132-141. http://dx.doi.org/10.1016/j.compedu.2009.01.008
*Shamir, A., Zion, M., & Spector_Levi, O. (2008). Peer tutoring, metacognitive processes and multimedia problem-based learning: The effect of mediation training on critical thinking. *Journal of Science Education and Technology, 17*, 384-398. http://dx.doi.org/10.1007/s10956-008-9108-4
*Shearer, J. E. (1981). *A comparison of individualized and group-oriented reading methods related to critical thinking* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8106340)
*Shepherd, N. G. (1998). *The probe method: A problem-based learning model's affect on critical thinking skills of fourth and fifth grade social studies students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9826025)
*Shin, S.-R. (2003). *The effects of a metacognitive art criticism teaching strategy that incorporates computer technology on critical thinking skill and art critiquing ability* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3072012)
*Sidney, A. H. (1989). *The effects of the inquiry method of teaching science on critical thinking skills, achievement and attitude toward science* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8917080)
*Slaughter, D. S., Brown, D. S., Gardner, D. L., & Perritt, L. J. (1989). Improving physical therapy students' clinical problem-solving skills: an analytical questioning model. *Physical Therapy, 69*(6), 441-447.
*Slavik, S. J. (1996). *An examination of the effects of selected disciplinary art teaching strategies on the cognitive development of selected sixth grade students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9544331)
*Smith, B. (1974). *Critically reading for propaganda techniques in grade six.*

(Unpublished master's thesis). Rutgers University, New Brunswick, NJ.
*****Solon, T. (2001). Improving critical thinking in an introductory psychology course. *Michigan Community College Journal: Research and Practice, 7*(2), 73-80.
*****Solon, T. (2007). Generic critical thinking infusion and course content learning in introductory psychology. *Journal of Instructional Psychology, 34*, 95-109.
*****Sondel, H. (2009). *The effects of curricular programs on aspects of critical thinking as applied to writing* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3364889)
*****Stambaugh, T. (2007). *Effects of the Jacob's Ladder Reading Comprehension Program on reading comprehension and critical thinking skills of third, fourth, and fifth grade students in rural schools* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3254406)
*****Starr, R. J. (1972). Structured oral inquiry improves thinking. *American Biology Teacher, 34*, 408-409.
*****Stoiber, K. C. (1991). The effect of technical and reflective preservice instruction on pedagogical reasoning and problem solving. *Journal of Teacher Education, 42*, 131-139. http://dx.doi.org/10.1177/002248719104200206
*****Suksringarm, P. (1976). *An experimental study comparing the effects of BSCS and traditional biology on achievement, understanding of science, critical thinking ability, and attitude toward science of the first year students at the Sakon Nakorn Teachers' College, Thailand* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7624805)
*****Sullivan-Mann, J., Perron, C. A., & Fellner, A. N. (2009). The effects of simulation on nursing students' critical thinking scores: A quantitative study. *Newborn and Infant Nursing Reviews, 9*, 111-116. http://dx.doi.org/10.1053/j.nainr.2009.03.006
*****Sungur, S., & Tekkaya, C. (2006). Effects of problem-based learning and traditional instruction on self-regulated learning. *Journal of Educational Research, 99*, 307-317. http://dx.doi.org/10.3200/JOER.99.5.307-320
*****Sykes, M. S. (1991). *The effects of instructional interventions for at-risk secondary students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9119219)
*****Szabo, Z., & Schwartz, J. (2009). Learning methods for teacher education: Blackboard discussions produce deep learning. In G. Siemens & C. Fulford (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2009* (pp. 2323-2334). Chesapeake, VA: AACE. Retrieved from
*****Tabbada, E. V. (1988). *Exploring the development of critical thinking in home economics* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8803216)
*****Tarkington, S. A. (1989). *Improving critical thinking skills using Paideia seminars in a seventh grade literature curriculum* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8912761)
*****Teagle, T. E. (1986). *The Socratic method of teaching: Its effect on the development of critical thinking skills of upper grade elementary school students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8617404)
*****Teele, L. (2006). *The impact of integrated study skills and critical thinking on student achievement* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3226249)

*****Tell, M. P. (1989). *The impact of a thinking skills curriculum on selected cognitive processes of intermediate grade students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8818721)

*****Thomas, J., Coppola, J. F., & Thomas, B. A. (2002). Smart e-classrooms, traditional classrooms and critical thinking. In G. Richards (Ed.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2002* (pp. 2288-2291). Chesapeake, VA: AACE. Retrieved from http://www.editlib.org/p/9750

*****Thomas, J. W. (1975, March-April). *The effects of instruction on the critical thinking abilities of middle-school-age children.* Paper presented at the annual meeting of the American Educational Research Association, Washington, D.C. (ERIC Document Reproduction Service No. ED105420)

*****Thompson, J. (2009). *To question or not to question: The effects of two teaching approaches on students' thinking dispositions, critical thinking skills, and course grades in a critical thinking course* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3355859)

*****Thompson, S. S. (1995). *Techniques for assessing general education outcomes: the natural science core program at the University of Denver (Colorado)* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9532955)

*****Thomson, S. (2009). *Effects of Primary Education Thinking Skills (TM) on achievement: Mixed-method evaluative study* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3401313)

*****Tilson, E. R. (1987). *The effect of computer enhanced skill training in critical thinking and cognitive monitoring on learning and development of critical thinking in an undergraduate family nursing course (Georgia)* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8706902)

*****Tiwari, A., Lai, P., So, M., & Yuen, K. (2006). A comparison of the effects of problem-based learning and lecturing on the development of students' critical thinking. *Medical Education, 40*, 547-554. http://dx.doi.org/10.1111/j.1365-2929.2006.02481.x

*****Topping, K. J., & Trickey, S. (2007). Collaborative philosophical inquiry for schoolchildren: Cognitive gains at 2-year follow-up. *British Journal of Educational Psychology, 77*, 787-796. http://dx.doi.org/10.1348/000709907X193032

*****Unan, V. S. S. (1979). *A comparison of the effects of BSCS and traditional biological science on achievement, critical thinking ability, and attitude toward science, scientists, and scientific careers of female slow learners in an Egyptian preparatory school* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7902661)

*****VanderStoep, S. W., & Shaughnessy, J. J. (1999). Taking a course in research methods improves reasoning about real-life events. In M. Ware & C. L. Brewer (Eds.), *Handbook for teaching statistics and research methods* (2 ed., pp. 242-244). Mahwah, NJ: Lawrence Erlbaum.

*****Vanlengen, C. A. (1989). *Effect of computer programming instruction on the problem solving ability of college-level introductory computer students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8907819)

*****VanTassel-Baska, J., Bracken, B., Feng, A., & Brown, E. (2009). A longitudinal

study of enhancing critical thinking and reading comprehension in title I classrooms. *Journal for the Education of the Gifted, 33*(1), 7-37.

*****Vasu, E. S., & Tyler, D. K. (1997). A comparison of the critical thinking skills and spatial ability of fifth grade children using simulation software or logo. *Journal of Computing in Childhood Education, 8*, 345-363.

*****Velde, B. P., Wittman, P. P., & Vos, P. (2006). Development of critical thinking in occupational therapy students. *Occupational Therapy International, 13*, 49-60. http://dx.doi.org/10.1002/oti.20

*****Waldron, J. M. (1992). *The effects of two instructional strategies for critical thinking-writing instruction on high ability middle school students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9231949)

*****Wallace, S. P. (1993). *A study of argumentative/persuasive writing related to a model of critical thinking in grades nine and eleven* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9239095)

*****Wallen, A., Reid, T., & Haubrick, V. (1963). The outcomes for critical thinking. *Journal of Educational Research, 56*, 529-534.

*****Wang, Y. (2009). Using discussion forum to enhance students' critical thinking ability through the Blackboard learning system. In T. Bastiaens et al. (Ed.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2009* (pp. 3243-3252). Chesapeake, VA: AACE.

*****Weast, D. (1996). Alternative teaching strategies: The case for critical thinking. *Teaching Sociology, 24*, 189-194.

*****Wells, L. A. (2002). Philosophy and psychiatry - A new curriculum for child and adolescent psychiatry. *Academic Psychiatry, 26*, 257-261. Retrieved from http://ap.psychiatryonline.org/cgi/reprint/26/4/257

*****Whalen, S. (1991). Intercollegiate debate as a cocurricular activity: Effects on critical thinking. In D. Parson (Ed.), *Argument in controversy: Proceedings of the 7th SCA/AFA Conference on Argumentation* (pp. 391-397). Annandale, VA: Speech Communication Association.

*****Whitten, M. M. (1971). *Changes in understanding of the nature of science and in critical thinking resulting from two different physical science courses* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 7122549)

*****Whyte, D. T. (1999). *The effect of an educational unit on the critical thinking skills of social work students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9841778)

*****Williams, R. L., Oliver, R., & Stockdale, S. (2004). Psychological versus generic critical thinking as predictors and outcome measures in a large undergraduate human development course. *Journal of General Education, 53*, 37-58.

*****Winocur, S. M. (1981). *The impact of a program of critical thinking on reading comprehension remediation and critical thinking of middle and high school students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8116378)

*****Wolf, W., King, M. L., & Huck, C. S. (1968). Teaching critical reading to elementary school children. *Reading Research Quarterly, 3*, 435-498.

*****Wright, A. (1991). *Critical thinking: Definition, measurement and development in vocational students in further education* (Doctoral dissertation). Available from

ProQuest Dissertations and Theses database. (UMI No. DX93650)
*Wu, W., & Yang, Y. C. (2009). Using a Multimedia Storytelling to Improve Students' Learning Performance. In G. Siemens & C. Fulford (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2009* (pp. 3159-3166). Chesapeake, VA: AACE. Retrieved from
*Xie, Y., Ke, F., & Sharma, P. (2008). The effect of peer feedback for blogging on college students' reflective learning processes. *Internet & Higher Education 11*, 18-25 http://dx.doi.org/10.1016/j.iheduc.2007.11.001
*Yalcin, B. M., Karahan, T. F., Karadenizli, D., & Sahin, E. M. (2006). Short-term effects of problem-based learning curriculum on students' self-directed skills development. *Croatian Medical Journal, 47*(3), 491-498. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2080434/pdf/CroatMedJ_47_0491.pdf
*Yang, S. C., & Chung, T.-Y. (2009). Experimental study of teaching critical thinking in civic education in Taiwanese junior high school. *British Journal of Educational Psychology, 79*, 29-55. http://dx.doi.org/10.1348/000709907X238771
*Yang, Y., & Chou, H. (2005). E-Learning: Learning How to Think through Interaction via Computer Conferencing. In P. Kommers & G. Richards (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2005* (pp. 886-890). Chesapeake, VA: AACE. Retrieved from
*Yang, Y.-T. C. (2004). *Use of structured web-based bulletin board discussions with Socratic questioning to enhance students' critical thinking skills in distance education* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3105051)
*Yang, Y.-T. C., & Chou, H.-A. (2008). Beyond critical thinking skills: Investigating the relationship between critical thinking skills and dispositions through different online instructional strategies. *British Journal of Educational Technology 39*, 666-684 http://dx.doi.org/10.1111/j.1467-8535.2007.00767.x
*Yang, Y.-T. C., Newby, T., & Bill, R. (2008). Facilitating interactions through structured web-based bulletin boards: A quasi-experimental study on promoting learners' critical thinking skills. *Computers & Education 50*, 1572-1585 http://dx.doi.org/10.1016/j.compedu.2007.04.006
*Yang, Y. C., & Li, L. (2009). Training pre-service teachers in how to teach critical thinking to middle school students. In T. Bastiaens et al. (Ed.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2009* (pp. 2060-2069). Chesapeake, VA: AACE. Retrieved from
*Yarema, C. H. (1996). *Effects of use of a context approach to calculus on students' critical thinking abilities, content acquisition, and attitudes toward mathematics* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9600133)
*Yeh, Y. C. (2009). Integrating e-learning into the direct-instruction model to enhance the effectiveness of critical-thinking instruction. *Instructional Science, 37*, 185-203. http://dx.doi.org/10.1007/s11251-007-9048-z
*Yuan, H., Kunaviktikul, W., Klunklin, A., & Williams, B. A. (2008). Improvement of nursing students' critical thinking skills through problem-based learning in the People's Republic of China: A quasi-experimental study. *Nursing & Health Sciences 10*, 70-76 http://dx.doi.org/10.1111/j.1442-2018.2007.00373.x
*Yuill, N. C. W. (1992). *The effects of computerized problem-solving simulations on critical thinking skills of baccalaureate nursing students* (Doctoral dissertation).

Available from ProQuest Dissertations and Theses database. (UMI No. 9203063)
*****Zohar, A., & Tamir, P. (1993). Incorporating critical thinking into a regular high school biology curriculum. *School Science and Mathematics*, *93*, 136-140.
*****Zohar, A., Weinberger, Y., & Tamir, P. (1994). The effect of the biology critical thinking project on the development of critical thinking. *Journal of Research in Science Teaching*, *31*, 183-196. http://dx.doi.org/10.1002/tea.3660310208

5. Avis pédagogiques

- Cook, J. (2017). How to Effectively Debunk Myths About Aging and Other Misconceptions. *Public Policy & Aging Report*, 27(1), 13-17.
- Maurage, P., Heeren, A., & Pesenti, M. (2013). Does chocolate consumption really boost Nobel award chances? The peril of over-interpreting correlations in health studies. *The Journal of Nutrition*, 143(6), 931-933.
- Messerli FH, *Chocolate consumption, cognitive function, and Nobel laureates*. N Engl J Med. 2012; 367 :1562–4.
- Matute, H., Blanco, F., Yarritu, I., Díaz-Lago, M., Vadillo, M. A., & Barberia, I. (2015). Illusions of causality: how they bias our everyday thinking and how they could be reduced. *Frontiers in Psychology*, *6*, 888.