



HAL
open science

Mécanismes quantitatifs et qualitatifs de la variabilité dans les tâches cognitives de haut niveau

Corentin Gonthier

► **To cite this version:**

Corentin Gonthier. Mécanismes quantitatifs et qualitatifs de la variabilité dans les tâches cognitives de haut niveau. Psychology. Université Rennes 2, 2020. tel-03289142

HAL Id: tel-03289142

<https://hal.science/tel-03289142>

Submitted on 16 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Rennes 2

Habilitation à Diriger des Recherches

Synthèse des travaux scientifiques présentée à l'appui de l'HDR

Mécanismes quantitatifs et qualitatifs de la variabilité dans les tâches cognitives de haut niveau

Corentin Gonthier

Département de Psychologie – Université Rennes 2

Laboratoire de Psychologie : Cognition, Comportement, Communication (EA 1285)

Soutenue publiquement le 08/12/2020

Devant le jury composé de :

- Pr. Maud Besançon – Université Rennes 2 (référente)
- Pr. Agnès Blaye – Université Aix-Marseille (présidente)
- Pr. Valérie Camos – Université de Fribourg (rapporteur)
- Pr. Jacques Grégoire – Université Catholique de Louvain (rapporteur)
- Pr. Jacques Juhel – Université Rennes 2 (rapporteur)

Mr. President, we are not "employees of the university".

We are the university.

– Université de Columbia, c. 1950

Attribué à Isidor I. Rabi, interrompant un discours de Dwight D. Eisenhower

I'm being quoted to introduce something,

*but I have no idea what it is
and certainly don't endorse it.*

– Randall Munroe, *xkcd* 1942

REMERCIEMENTS

Si cette section "remerciements" paraît dithyrambique, c'est surtout parce que j'ai eu beaucoup de chance dans mon parcours et que j'y ai essentiellement côtoyé des personnes qui méritent des compliments.

Mes premiers remerciements s'adressent au jury qui a eu la générosité de donner de son temps pour évaluer ce travail d'habilitation : Maud Besançon, Agnès Blaye, Valérie Camos, Jacques Grégoire, et Jacques Juhel - tous des chercheurs de la plus grande qualité et qui m'honorent en me prêtant leur expertise, et je ne le dis pas uniquement parce qu'ils lisent ces lignes et vont devoir m'évaluer à brève échéance.

Mes seconds remerciements s'adressent à mes collègues du département de psychologie de l'Université Rennes 2, qui ont eu la gentillesse de m'accueillir après ma thèse de doctorat et la très grande gentillesse de ne pas trop s'en plaindre. Un grand merci en particulier à Séverine, dont la sincérité et l'enthousiasme scientifique lui auront permis de mériter comme une évidence son diplôme de "meilleure collègue de bureau du monde", à accrocher à côté de son HDR. Merci également à Maud, Thierry, Géraldine et Jacques, qui ont fait de l'équipe de psychologie différentielle de Rennes 2 le cadre le plus familial, le plus soutenant et le plus agréable dont un différentialiste puisse rêver aujourd'hui.

En plus d'être une personne et une enseignant-chercheur intrinsèquement fantastique, Maud référentise mon HDR, ce qui la rendrait encore plus fantastique si toutefois c'était possible... ce dont je doute fortement : Maud maintient une activité de recherche et d'enseignement complète, dirige notre laboratoire ainsi qu'un master, siège au CAC, à la CFVU, au conseil d'UFR et à l'association de parents d'élèves pour ses trois enfants, et m'a envoyé sa relecture finale de ce document à 07h56 du matin. Il est donc évident que le mérite de la relecture revient à Maud et que les erreurs me reviennent à moi.

Parmi mes collègues du département de psychologie de Rennes 2, je souhaite remercier plus particulièrement, pour nos échanges et/ou pour leur accueil : Éric, Virginie Dodeler (nous attendons avec impatience la publication de son échelle d'évaluation de la qualité des

services institutionnels universitaires – Item 1 : "Faut tout cramer et virer tout le monde"), Zdenka, Fanny, Nathalie, Olivier, Audrey et Alessandro.

Sur le plan scientifique, je dois une bonne part de mon dossier de publications aux collaborations scientifiques proposées par mes collègues : en Bretagne, Éric et Séverine, Nathalie Marec-Breton et (Em)Manu(elle), Maud, Zdenka, Fanny et Nathalie Bonneton-Botté, Florence Gaillard et Armel Cretual ; en France, Virginie Bagneux, Pascal Hot, Morgane Metral, Vera Tsenova, Alexandre Aubry et Béatrice Bourdin, Agnès Blaye, et Lucia Romo ; et ailleurs, l'équipe de Todd Braver, Julie Bugg, l'équipe d'Andrew Conway, et l'équipe d'Ettore Ambrosini.

Agnès en particulier a eu la gentillesse de m'inviter à collaborer sur les aspects développementaux du contrôle cognitif et m'a donné le courage de m'attaquer à la thématique du développement cognitif en général ; sans son implication, j'aurais certainement clôturé cette thématique de recherche il y a plusieurs années.

Ceux qui les connaissent bien s'apercevront sans difficulté que j'ai volé à mon directeur de thèse Jean-Luc Roulin, ainsi qu'à Jacques Juhel, leurs meilleures idées pour les mettre dans cette HDR, et plus généralement dans mon activité de recherche. Soi-disant qu'on doit se détacher progressivement de l'influence intellectuelle de son directeur de thèse. Je n'ai visiblement pas encore atteint ce stade de ma carrière. Je commence à douter de l'atteindre un jour et à prendre conscience, progressivement, de l'horrible vérité : je dirai probablement la même chose dans ma présentation de départ à la retraite.

Noémyle, qui a ajouté à son doctorat en psychologie différentielle une Habilitation à Diriger ma Vie (HDV), mais en use avec discernement. En général¹.

¹ En fait, je n'ai jamais voulu faire de la recherche : ça, c'était l'idée de Noémyle (qui a toujours fourni l'essentiel de la force motrice au sein de notre couple). Au second jour de la Licence 1, elle proclama : "moi, je ferai une thèse !". Dont acte. Je l'ai suivie avec plus d'incertitude, surtout par dépit face à l'ennui écrasant auquel m'ont confronté les bilans Alzheimer lors d'un stage clinique.

Ensuite, Noémyle a décidé de travailler sur les liens entre contrôle cognitif, mémoire de travail et intelligence. Ça avait l'air d'être une bonne idée, alors je me suis greffé sur son projet.

Après quoi, Noémyle a décidé que les stratégies n'avaient pas reçu suffisamment d'attention dans la littérature sur la cognition de haut niveau, et a choisi d'en faire le sujet de sa thèse. Quand elle a eu fini avec ce thème, je me suis permis de le récupérer.

À ce stade du propos, le lecteur attentif ayant astucieusement remarqué que cette habilitation à diriger des recherches porte sur le *contrôle cognitif, la mémoire de travail et l'intelligence* et notamment sur les variations qualitatives de *stratégies* se demandera ce que diable j'ai bien pu choisir par moi-même dans mon parcours professionnel. Très simple : j'ai choisi Noémyle.

SOMMAIRE

REMERCIEMENTS	- 2 -
SOMMAIRE	- 4 -
INTRODUCTION	- 6 -
1. Ma vie, mon œuvre	- 6 -
2. Ma posture épistémologique	- 8 -
CHAPITRE 1. CONTRÔLE COGNITIF.....	- 17 -
1. Revue de littérature	- 17 -
2. Évaluer les mécanismes de contrôle	- 27 -
2.1. Le modèle DMC à travers l'AX-CPT	- 27 -
3.2. Le modèle DMC à travers les effets de proportion de congruence.....	- 32 -
3. Étudier le développement du contrôle cognitif	- 40 -
3.1. Le développement du contrôle à travers l'AX-CPT	- 40 -
3.2. Le développement du contrôle à travers les effets de proportion de congruence..	- 43 -
4. Induire l'usage d'un mécanisme de contrôle	- 47 -
4.1. Induire le contrôle proactif et réactif chez l'adulte	- 47 -
4.2. Induire le contrôle proactif chez l'enfant ?	- 50 -
5. Perspectives de recherche	- 52 -
5.1. Axe 1 : Trois enjeux pour le modèle DMC.....	- 52 -
5.2. Axe 2 : Dynamiques du contrôle dans les tâches de conflit et les effets de proportion de congruence.....	- 57 -
5.3. Axe 3 : Variabilité du contrôle cognitif.....	- 60 -
6. Publications et communications associées.....	- 62 -
CHAPITRE 2. MÉMOIRE DE TRAVAIL	- 64 -

1. Revue de littérature	- 64 -
2. Évaluer la capacité en mémoire de travail	- 74 -
2.1. En tenant compte de son caractère domaine-spécifique	- 74 -
2.2. En tenant compte de la longueur de la tâche et du niveau d'aptitude	- 79 -
3. Mécanismes de la performance en mémoire de travail	- 83 -
3.1. Étapes du processus mnésique	- 83 -
3.2. Le rôle du contrôle cognitif	- 86 -
3.3. Le rôle des stratégies.....	- 92 -
4. Mémoire de travail et cognition de haut niveau	- 95 -
4. Perspectives de recherche	- 102 -
4.1. Axe 1 : Outils d'évaluation de la capacité en mémoire de travail.....	- 102 -
4.2. Axe 2 : Déterminants stratégiques de la performance en mémoire de travail	- 110 -
5. Publications et communications associées	- 112 -
CHAPITRE 3. INTELLIGENCE FLUIDE.....	- 115 -
1. Revue de littérature	- 115 -
2. Le rôle des stratégies dans les tests d'intelligence	- 128 -
3. Le rôle du temps dans les tests d'intelligence	- 137 -
4. Le rôle des connaissances dans les tests d'intelligence	- 143 -
5. Perspectives de recherche	- 152 -
5.1. Sur le rôle des stratégies dans les tests d'intelligence	- 152 -
5.2. Sur les autres déterminants de la performance	- 155 -
5.3. Vers une architecture causale du raisonnement (CAR)	- 157 -
6. Publications et communications associées	- 164 -
SYNTHÈSE ET INDEX THÉMATIQUE	- 166 -
BIBLIOGRAPHIE.....	- 171 -

INTRODUCTION

1. Ma vie, mon œuvre

Je suis arrivé à la psychologie différentielle par la clinique. Ma formation initiale à l'Université de Savoie était celle d'un neuropsychologue, et je me destinais à une activité de psychologue spécialisé en psychopathologie cognitive. La réalité de l'activité clinique, à travers sa composante évaluative, m'a vite confronté à une problématique centrale : pourquoi mon patient a-t-il chuté sur ce test ? Face à une faible performance, on cherche à répondre à la question d'un éventuel déficit. Pour cela, il faut bien sûr s'interroger sur les déterminants de la performance – quels sont ces déterminants, et parmi ces déterminants, lesquels sont susceptibles d'être à l'origine de la chute observée ? S'agit-il vraiment d'un trouble dysexécutif, ou bien de fatigue, ou d'une faible motivation ? Interpréter un score impose naturellement une réflexion causale, intégrant la pluralité des mécanismes à l'origine de la variabilité.

Ce n'est pas un hasard si mon premier terrain de stage en neuropsychologie, l'unité de neurologie vasculaire du CHU de Grenoble, a aussi été celui de ma première expérience de recherche dans le champ de la cognition de haut niveau. J'ai mené ce premier travail en Master 1 avec Noémie Thomassin, sous la direction de Pascal Hot et Bernadette Naëgelé. Notre questionnement consistait à identifier l'origine des déficits observés à la suite d'un accident vasculaire cérébral dans les domaines de la mémoire de travail et du fonctionnement exécutif, même en cas de lésions situées en-dehors des régions exécutives telles que le cortex préfrontal. Pour cela, nous avons dû réfléchir à la diversité des déterminants de la performance exécutive pouvant être affectés par une lésion cérébrale et donc jouer un rôle causal dans ces déficits (Gonthier, Thomassin, Naëgelé, Hot, & Detante, 2010).

Nous sommes arrivés à la conclusion que les patients atteints d'un accident vasculaire cérébral devenaient significativement *meilleurs* que les sujets contrôles dans un certain nombre de tests exécutifs. Ce résultat, que nous avons attribué à une différence d'engagement dans la tâche entre sujets contrôles et cérébrolésés (passation chez soi en

pantoufles pour faire plaisir à un étudiant, versus évaluation d'un déficit post-accident vasculaire cérébral à l'hôpital par un professionnel en blouse blanche), n'a fait que me pousser dans la direction d'un scepticisme vis-à-vis du score total et d'une réflexion sur les multiples mécanismes de la variabilité.

Par la suite, mon activité de recherche a été guidée par cette question de la multi-détermination : tenir compte de la pluralité des mécanismes causals amenant à déterminer le score total, pour mieux interpréter la variabilité. À travers cette question, la psychologie clinique apparaît évidemment comme une psychologie différentielle – ou la psychologie différentielle comme une psychologie clinique : qu'est-ce qui fonde la spécificité de mon sujet, quels mécanismes causals sont à l'origine de son pattern particulier de performance, quelle est la nature du fonctionnement ou du dysfonctionnement qui lui est spécifique ? S'il m'a fallu attendre 2017 et ma troisième année d'exercice en tant que maître de conférences pour finalement suivre un Master 2 professionnel, obtenir mon titre de psychologue clinicien et compléter cette partie de mon parcours, j'ai toujours vu une continuité entre ces deux aspects d'un même questionnement portant sur l'origine des scores et de leur variabilité.

À ce niveau, notre discipline semble parfois souffrir d'un léger trouble dissociatif. Tous les neuropsychologues apprennent comme une évidence qu'un score faible au Trail Making Test (TMT) n'est pas forcément le symptôme d'un trouble de la flexibilité mentale, une faible performance pouvant tout aussi bien s'expliquer par un ralentissement psychomoteur ou même une perte d'automatisation de l'ordre alphabétique chez la personne âgée. En parallèle, la littérature présente des piles vertigineuses d'articles dans lesquels le score au TMT est considéré comme *la flexibilité mentale* et utilisé en l'état pour prédire d'autres indicateurs (démarche que mon directeur de thèse Jean-Luc Roulin qualifiait de "pêche aux corrélations"), sans jamais mener de réflexion sur les multiples facteurs qui avaient déterminé la variabilité du score en premier lieu et qui pourraient, donc, être à l'origine des corrélations observées.

Mon questionnement central concerne donc l'interprétation de la performance dans les tâches cognitives de haut niveau, et donc de la variabilité de la performance sous toutes ses formes (inter ou intra-individuelle, adulte ou développementale...). Ce questionnement m'a amené à m'intéresser en particulier aux tâches de contrôle cognitif (Chapitre 1), de mémoire de travail (Chapitre 2), et d'intelligence fluide (Chapitre 3).

2. Ma posture épistémologique

Les psychologues conçoivent surtout la performance sur les tâches cognitives de haut niveau selon une posture *réaliste* : une tâche mesure une entité qui existe *réellement*. Cette entité est bien définie, bien délimitée par rapport à d'autres aspects de la cognition ; elle a une existence concrète, indépendamment de la façon dont on la définit et dont on la mesure (cette existence concrète étant souvent argumentée en s'appuyant sur l'alibi d'un substrat neuronal qui lui serait propre) ; elle est essentiellement unitaire. Une tâche mesure "le contrôle cognitif", "la mémoire de travail", "l'intelligence" : ces trois concepts sont vus comme des aptitudes cognitives qui existent intrinsèquement et qui, de ce fait, déterminent directement le score total sur les tâches qui en relèvent.

La posture réaliste est bien pratique pour travailler sur les différences inter-individuelles : elle permet une certaine croyance en la réalité des scores que l'on collecte, et une croyance dans le rôle causal que l'aptitude reflétée par ces scores peut avoir pour l'individu (il a réussi la tâche parce qu'il est intelligent ; cette intelligence causera sa réussite dans d'autres domaines). Comme beaucoup d'auteurs, je crois que la posture réaliste est une simplification indéfendable (e.g. Borsboom et al., 2003, 2009; Kovacs & Conway, 2016), et que la mesure en psychologie n'autorise aucune conclusion directe sur l'existence de l'entité qui la provoque.

L'argument central en faveur de la position réaliste est toujours l'observation que les différentes tâches censées évaluer une entité se regroupent en un même facteur statistique (mémoire de travail : Kane et al., 2004 ; contrôle cognitif : Miyake et al., 2000 ; intelligence : Spearman, 1904) ; mais l'analyse factorielle ne fait qu'analyser les corrélations entre tests, sans pouvoir évidemment donner d'indications sur l'ontologie de la ou des entités qui en sont à l'origine (Burt, 1940 ; Jensen, 1998 ; Spearman, 1927). Les analyses factorielles semblent générer nettement plus de confusion à ce niveau que les analyses de corrélation, en dépit de leur équivalence statistique, peut-être parce que la foi aveugle que les psychologues accordent à leurs concepts croît avec la complexité des analyses qui permettent de les étudier (on retrouve souvent un phénomène similaire dans le cas des analyses en équations structurales ou des analyses dites "en pistes causales").

Plus généralement la mesure, qui situe les individus relativement les uns aux autres, est un niveau de description *inter-individuel* qui ne fournit aucune information sur le

processus qui a donné lieu à ce classement relatif au niveau *individuel* (Borsboom, 2003 ; Molenaar, 2004 ; Lamiell, 1981, Mischel & Shoda, 1995). Croire en la réalité de l'entité "intelligence", c'est donc partir de la variation observée entre les individus pour la généraliser abusivement à l'existence d'une entité causale qui existe au niveau individuel. Cette généralisation du niveau inter-individuel au niveau individuel ne peut être valable, sauf à supposer que la structure de la variation inter-individuelle (la variation *entre les individus* sur une même tâche se résume de façon homogène à leur facteur général de l'intelligence) est un reflet exact du processus au niveau individuel (la variation *d'un même individu entre différentes tâches* se résume de façon homogène à son facteur d'intelligence). On peut qualifier ce postulat d'homologie (Hamaker et al., 2005).

Un parfait respect du postulat d'homologie requiert plusieurs conditions, de la moins stricte à la plus stricte (Borsboom, 2009) :

- homogénéité locale (le modèle de mesure moyen est valable pour chaque sujet : le nombre de facteurs et plus généralement les corrélations entre tâches sont les mêmes au niveau intra-individuel qu'au niveau inter-individuel ; autrement dit, deux sujets qui présentent le même niveau sur la variable latente causale ont le même pattern prédit de scores sur un ensemble de subtests, et il n'existe pas de sous-population qui présente un pattern différent ; Borsboom, 2009 ; Ellis & van den Wollenberg, 1993)
- invariance stricte de la mesure à travers les individus (une différence de score moyenne entre deux individus ne peut être causée que par la variable latente causale)
- ergodicité (la structure de la variation entre les individus est la même que la structure de la variation au sein d'un individu, ce qui implique non seulement que les corrélations entre tâches soient stables à travers la trajectoire développementale d'un individu et que toutes les trajectoires développementales soient parallèles, mais aussi qu'il n'existe pas de différences inter-individuelles en moyenne ; Molenaar, 2004 ; Lautrey, 2003).

Le postulat d'homologie apparaît franchement intenable dans le cas de la cognition de haut niveau : même la condition la plus simple d'homogénéité locale n'est pas satisfaite. Elle implique que tous les sujets avec le même niveau d'aptitude doivent avoir le même pattern de réponses, et qu'il n'existe pas de sous-populations avec des profils différents : or, cette proposition est immédiatement invalidée par l'existence de profils pathologiques qui

présentent des forces et des faiblesses spécifiques sur différents subtests d'intelligence (dyslexie, dyspraxie par exemple). Argumenter en faveur de l'homologie négligerait la littérature abondante qui insiste sur la pluralité des mécanismes, à la fois quantitatifs et qualitatifs, à l'origine de la performance : croire que la performance individuelle à travers l'ensemble des tests est parfaitement résumée par un unique facteur général, c'est supposer qu'il n'existe pas d'autre déterminant transversal de la performance sur les tests d'intelligence que l'entité "intelligence", qu'il n'y a qu'une façon possible de résoudre une tâche, et qu'il n'y a qu'une façon d'obtenir un score d'intelligence donné.

Je crois que ce raisonnement amène logiquement à l'adoption d'un point de vue *constructiviste* (Borsboom, 2003) : contrôle cognitif, mémoire de travail et intelligence sont des constructions des chercheurs en psychologie, qui ne renvoient pas à des entités discrètes qui auraient une réalité au niveau de l'individu. Ces construits n'existent pas au-delà de la mesure et de la conceptualisation que l'on en fait : ce sont des outils de convenance référentielle, utiles pour désigner des phénomènes qui existent bel et bien sur le plan empirique, mais qui ne doivent pas être réifiés en supposant que ces phénomènes sont créés par une entité distinctive. La distinction entre réalisme et constructivisme n'est pas qu'un débat philosophique (pour ne pas dire théologique) : elle a un certain nombre de conséquences pratiques pour l'interprétation de la variabilité. Mon interprétation de ces conséquences est la suivante :

- 1) La posture **réaliste** considère la mesure comme la transcription directe (à l'erreur de mesure près) d'une entité sous-jacente : le score sur une tâche de contrôle cognitif est *causé par l'entité "contrôle cognitif"*, qui constitue une aptitude à part entière. Le contrôle cognitif est une *cause* d'autres mesures. La posture **constructiviste** considère la mesure comme émergeant du fonctionnement d'un ensemble de mécanismes dans un contexte donné : le score de contrôle cognitif est une *propriété émergente d'une constellation de processus cognitifs* en interaction. Le contrôle cognitif est une *conséquence* d'autres mesures.
- 2) La posture **réaliste** décrit le score sur un test de mémoire de travail comme *la mémoire de travail*, raccourci justifié par le lien causal direct entre le score et l'entité qui le provoque. La posture **constructiviste** préfère parler de *la performance sur une tâche de*

mémoire de travail, cette périphrase permettant d'éviter l'hypothèse implicite d'une entité causale.

- 3) La posture **réaliste** modélise "l'intelligence" de façon *réflective* : la variable latente (l'intelligence) est *reflétée* par des variables manifestes (le score aux tests d'intelligence) ; ces variables manifestes sont affectées par une erreur de mesure (la performance peut ne pas refléter précisément la variable latente) ; des corrélations entre variables manifestes existent nécessairement et sont exclusivement causées par la variable latente. La posture **constructiviste** modélise "l'intelligence" de façon *formative* : la variable latente (l'intelligence) est *formée* à partir de variables manifestes (les différents processus en jeu dans un test d'intelligence) ; s'il doit y avoir une variance d'erreur, on la situe au niveau de la variable latente (la somme des processus que l'on a mesurés peut ne pas suffire à situer précisément la position sur la variable latente qui résume la performance finale) ; les corrélations entre variables manifestes peuvent être nulles ou non et ne sont pas particulièrement pertinentes pour le modèle.
- 4) La posture **réaliste** part du principe que le score est causé par une entité et dépend du degré d'efficacité *de cette entité*. Un score *correspond à une position précise* sur ce continuum d'efficacité de l'entité. La posture **constructiviste** part du principe que le score est causé par un système de processus en interaction et dépend du degré d'efficacité *de l'ensemble de ce système* (multi-détermination). Un score *ne correspond pas à une position précise* sur un continuum unique : un même score peut être obtenu par différentes combinaisons (multi-réalisabilité ; Borsboom, 2009) à la fois quantitatives - différents degrés d'efficacité sur les divers processus nécessaires à la performance, qui peuvent se compenser entre eux - et qualitatives - différentes approches stratégiques de la tâche. Une conséquence majeure de ce point est que le score total représente le mélange des influences de différents déterminants ; la corrélation entre le score total et une autre variable peut donc être causée par n'importe lequel ou lesquels de ces déterminants du score (e.g. Glymour, 1998).
- 5) La posture **réaliste** s'intéresse avant tout au niveau *inter-individuel* : puisque "l'intelligence" existe, le niveau d'analyse approprié est l'intelligence, et la façon dont le score d'intelligence corrèle avec d'autres variables ; il n'est pas particulièrement indispensable d'étudier les processus à l'origine de ce score. La posture **constructiviste**

s'intéresse au niveau *individuel* : puisque "l'intelligence" n'existe pas, le niveau d'analyse approprié est celui des différents mécanismes mis en jeu par l'individu pour réussir une tâche d'intelligence.

(Dans ce contexte, on oppose parfois étude des différences inter-individuelles et étude des différences intra-individuelles, ce qui est à mon avis injustifié : la posture réaliste peut aborder la variation intra-individuelle en la résumant par une variable exclusivement inter-individuelle, comme le "niveau de stabilité", sans faire d'hypothèses sur les mécanismes en jeu au niveau individuel ; tout comme la posture constructiviste peut aborder les mécanismes individuels de la performance sur une tâche en testant si cette performance est prédite au niveau inter-individuel par le niveau d'efficacité sur des mécanismes qui sont censés y contribuer, de façon à générer des hypothèses qui s'appliqueront ensuite au niveau intra-individuel. La distinction porte bien sur le niveau auquel les résultats sont interprétés - entité réaliste ou processus qui y contribuent -, plus que sur le type inter ou intra de variabilité étudiée sur le plan expérimental.)

6) Le niveau d'analyse inter-individuel qui découle de la posture **réaliste** s'oriente vers la *description des différences inter-individuelles* ("les tests d'intelligence corrèlent entre eux sous la forme d'un facteur général"), et vers une forme de *prédiction inter-individuelle* ("un sujet dont le score est X devrait, en moyenne, plutôt réussir la tâche Y, à condition que le sujet et la tâche fassent partie de la population idéale de sujets et de tâches auxquels mon modèle de mesure se destine"). En revanche, elle ne fournit pas d'éléments pour l'interprétation de la *causalité* : la variation inter-individuelle identifiée n'a pas forcément elle-même de valeur causale puisqu'elle n'est que la conséquence d'autres mécanismes fonctionnant au niveau individuel, mais non étudiés (Burt, 1940). La corrélation entre la mesure inter-individuelle de l'intelligence et une autre mesure est donc intrinsèquement suspecte car elle peut être causée par le rôle de n'importe lequel des mécanismes qui contribuent au score d'intelligence. Le niveau d'analyse individuel qui découle de la posture **constructiviste** privilégie la *description au niveau individuel*, c'est-à-dire construite à partir de mécanismes définis au niveau individuel ("ce sujet est décrit par sa motivation, sa stratégie, le nombre d'hypothèses qu'il génère..."), et la *prédiction au niveau individuel* ("étant donnés ses niveaux d'efficacité sur les différents processus A, B et C mis en jeu par les contraintes X, Y et Z de la tâche, ce sujet devrait

plutôt réussir la tâche") ; il autorise aussi l'interprétation de la *causalité* puisque l'unité d'analyse est bien constituée des processus individuels qui sont réellement à l'origine de la performance.

- 7) La posture **réaliste** est avant tout une approche nomothétique : elle s'intéresse à fournir des lois générales pour comprendre la nature des phénomènes en jeu aussi bien que leur variabilité inter-individuelle. La posture **constructiviste** est également une approche nomothétique puisqu'elle recherche la formulation de lois générales ; mais parce qu'elle se situe au niveau de description individuel et qu'elle analyse les mécanismes qui produisent l'échec ou la réussite pour un individu donné, l'approche constructiviste a l'ambition de fournir des éléments qui permettent l'approche idiographique, c'est-à-dire qui autorisent l'analyse des particularités d'un sujet donné en tant qu'elles découlent des lois générales que l'on a identifiées par ailleurs (voir Borsboom, 2003).
- 8) La posture **réaliste** va plutôt de pair avec une approche *structurelle* : on se concentre sur la description de différentes entités statiques à l'origine du phénomène ("*quoi ?*"). La posture **constructiviste** va plutôt de pair avec une approche *fonctionnelle* (l'étude du fonctionnement d'un système n'imposant pas de réalisme vis-à-vis des entités qui le composent) : puisque l'on part du principe qu'il n'existe pas forcément d'entités causales bien définies, on va plutôt s'intéresser à la description du fonctionnement du phénomène situé dans son contexte, aux étapes du processus qui mène à la performance, au rôle fonctionnel que servent ces différentes étapes ("*comment ?*"). Cette distinction entre structure et fonction n'est pas absolue, la plupart des modèles intègrent ces deux aspects, et il n'y a pas de recouvrement exact avec la posture épistémologique ; mais en ce qui concerne les trois construits dont je parlerai ici, les auteurs qui ont mis l'accent sur une approche fonctionnelle ont aussi plutôt adopté une posture constructiviste et vice versa.

Ces éléments de comparaison sont synthétisés dans le Tableau 1 ci-dessous.

Tableau 1. Proposition de synthèse des divergences entre posture réaliste et constructiviste

	Réaliste	Constructiviste
1) Qu'est-ce que le phénomène ?	Le produit d'une entité qui existe concrètement	La propriété émergente d'un système de processus en interaction
2) Comment désigner le phénomène ?	Sous le nom de l'entité supposée : "l'intelligence"	Comme la réussite sur une tâche : "la performance sur un test d'intelligence"
3) Comment modéliser le phénomène ?	Comme une variable réflexive : la variable latente existe et se reflète dans les mesures qu'elle détermine	Comme une variable formative : la variable latente n'existe que comme la somme d'un ensemble de déterminants
4) Que reflète le score à une tâche ?	L'efficacité de l'entité correspondante, qui correspond à une position précise sur un continuum unidimensionnel de l'efficacité	L'efficacité d'un système de processus dans son ensemble (multi-détermination), qui peut correspondre à plusieurs positions pour chaque processus (multi-réalisabilité)
5) À quel niveau doit s'étudier le phénomène ?	Surtout au niveau inter-individuel : puisque l'entité existe, c'est l'efficacité de cette entité qui compte	Surtout au niveau individuel : puisque l'entité n'existe pas, ce sont les mécanismes causals en jeu dans la tâche qui comptent
6) Que peut-on faire avec notre modèle du phénomène ?	A) Décrire l'organisation des différences inter-individuelles B) Prédire la performance moyenne au niveau inter-individuel, mais pas pour un individu donné	A) Décrire l'organisation des différences inter ou intra-individuelles B) Prédire la performance au niveau inter ou intra-individuel C) Expliquer les causes de la performance au niveau individuel
7) Quelles démarches sont possibles ?	Nomothétique : on cherche à extraire des lois générales qui s'appliquent en moyenne	Nomothétique, mais aussi idiographique : on peut chercher à interpréter la spécificité d'un sujet donné comme l'instantiation de l'ensemble des mécanismes généraux
8) Quel est le meilleur moyen d'aborder la description du phénomène ?	L'approche structurelle : on cherche à décrire les différentes entités qui existent et qui composent la structure de la cognition	L'approche fonctionnelle : on cherche à décrire par quel processus le phénomène émerge dans le cadre d'une tâche donnée, quels mécanismes contribuent à ce processus et quels rôles ils jouent

L'adoption d'une posture constructiviste m'a conduit à aborder la variabilité de la performance comme le résultat du fonctionnement d'un ensemble de mécanismes quantitatifs et qualitatifs, et à m'intéresser à l'étude de ces mécanismes et aux façons possibles d'en tenir compte. Dans le cadre de mon activité de recherche, j'ai donc cherché à étudier, modéliser, mesurer, ou contourner, les différents mécanismes déterminant la variabilité de la performance dans les tâches cognitives de haut niveau, dans leur pluralité. Les deux piliers de cette démarche sont d'une part la reconnaissance du caractère massivement multi-déterminé de la performance, et d'autre part l'intérêt pour la spécification des conditions de mise en œuvre des mécanismes qui contribuent de façon causale à la performance au niveau individuel, selon une approche plutôt fonctionnelle que structurelle. Ces deux préoccupations s'appliquent de façon transversale aux trois construits étudiés ici - contrôle cognitif, mémoire de travail et intelligence.

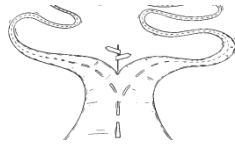
L'étude des mécanismes qualitatifs de la performance, en particulier, constitue un point d'intérêt central dans mes recherches, que j'ai décliné à ces trois thématiques (contrôle proactif ou réactif dans le cas des tâches de contrôle cognitif, stratégies de mémorisation dans le cas des tâches de mémoire de travail, stratégies de raisonnement dans le cas des tâches d'intelligence fluide). Il y a au moins trois raisons à cet intérêt : les stratégies ont été peu étudiées, et souvent sous l'angle d'un facteur de nuisance ; elles constituent un mécanisme causal de la performance facilement interprétable au niveau individuel ; et enfin un changement de stratégie peut modifier radicalement la nature de la tâche - à tel point que deux sujets utilisant des stratégies différentes réalisent en substance deux tâches différentes puisque le processus de traitement de l'information mis en jeu n'est pas le même. L'existence d'une variabilité dans les mécanismes qualitatifs de la performance constitue donc un argument central en faveur de la position constructiviste : des déterminants quantitatifs de la performance peuvent être négligés (avec un argument de type "ce n'est que de la variance d'erreur"), mais la variabilité qualitative remet fondamentalement en cause l'interprétation que l'on doit faire de la nature d'une épreuve.

Le présent document est peut-être un peu dense, ce dont je m'excuse par avance auprès de vous, cher lecteur ; son unité se trouve moins dans les objets d'étude que dans le type de questionnement qui me préoccupe. Ce questionnement central est celui du constructivisme, détaillé plus haut ; mais quelques thématiques plus précises reviennent de façon récurrente à travers les domaines étudiés. Un petit lexique de ces thématiques récurrentes est proposé ci-dessous, assorti de symboles ; dans le reste du document, je fais apparaître ce symbole à la marge lorsque je répète l'une de ces thématiques, comme une sorte de repère.

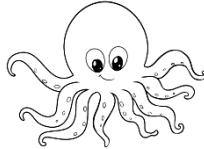
- Difficulté à étudier les mécanismes de la variabilité d'une entité considérée comme réaliste (donc relativement unitaire), et nécessité de s'intéresser à l'aspect fonctionnel, en examinant la façon dont l'activité cognitive est déployée par un participant sur un essai :



- Intérêt d'étudier l'existence de plusieurs mécanismes qualitatifs (notamment sous l'angle de différentes stratégies) pour comprendre le fonctionnement du sujet :



- Caractère massivement multi-déterminé de la performance sur une tâche complexe :



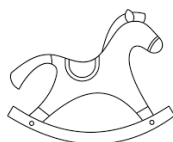
- Une conséquence de la multi-détermination : la corrélation entre deux construits ou mesures peut s'expliquer par le fait qu'ils partagent n'importe lequel de leurs déterminants, sans qu'il s'agisse forcément de l'objet d'intérêt ; si la tarte est bonne, cela peut s'expliquer par n'importe lequel de ses ingrédients causals :



- Utilisation de la démarche expérimentale-corrélationnelle (Cronbach, 1957) pour dépasser le problème précédent et vérifier quel déterminant causal est réellement à l'origine de la corrélation entre deux construits :



- Une autre conséquence de la multi-détermination : les frontières entre entités deviennent floues puisqu'elles partagent certains de leurs déterminants, ce qui implique que les construits se recouvrent et se comprennent mieux comme la propriété émergente de systèmes qui se chevauchent partiellement :



CHAPITRE 1. CONTRÔLE COGNITIF

Résumé

L'approche dominante du contrôle cognitif est plutôt réaliste et structurelle : elle consiste à postuler une pluralité de fonctions exécutives (classiquement inhibition, flexibilité et mise à jour), que l'on invoque ensuite comme explication à la variabilité observée sur différentes tâches. Une autre façon de penser le contrôle cognitif consiste à s'intéresser au plan fonctionnel, et à spécifier plus précisément le processus par lequel le contrôle cognitif est implémenté en réponse à un stimulus.

Mes propres travaux ont porté sur un modèle fonctionnel particulier, le modèle à deux mécanismes de contrôle (Dual Mechanisms of Control - DMC). Le modèle propose deux mécanismes possibles, proactif et réactif ; l'usage de l'un ou l'autre mécanisme est vu comme étant à l'origine de la variabilité des performances. Autour de ce modèle, j'ai travaillé sur trois points de questionnement principaux : 1) la mesure du contrôle proactif et réactif, à l'aide du paradigme AX-CPT et de tâches de conflit de type Stroop ; 2) le développement des deux mécanismes chez les enfants d'âge préscolaire ; 3) l'induction expérimentale de l'usage de l'un ou l'autre mécanisme.

1. Revue de littérature

Le "contrôle cognitif" désigne le fait de réguler son activité cognitive et son comportement de façon à atteindre un but particulier, dans une situation qui ne peut pas être résolue de façon automatique. Le contrôle cognitif se définit donc en creux : il n'est pas impliqué lorsque le contexte nécessite uniquement la mise en jeu automatique de processus cognitifs élémentaires (perception, mémoire...). Ainsi, la correction d'une copie d'examen de

Licence 1 pouvant être réalisée de façon purement automatique à l'aide de scripts pré-établis et d'actions perceptives et motrices élémentaires, ne nécessiterait pas de contrôle cognitif. Le contrôle cognitif est en revanche susceptible d'être recruté dès lors que le fonctionnement automatique de ces processus élémentaires ne permet pas à l'individu d'atteindre le but fixé, et qu'une régulation est nécessaire : le rôle du contrôle cognitif est souvent décrit à travers l'image d'un chef d'orchestre. Norman et Shallice (1986) proposent une liste utile des situations qui demandent un contrôle cognitif : celles qui nécessitent une planification ou une prise de décision complexe, dont le traitement automatique conduit à une erreur, qui sont nouvelles ou ne sont pas suffisamment maîtrisées pour autoriser un traitement automatique, qui sont difficiles ou dangereuses. On peut donc supposer que la lecture critique d'un manuscrit d'habilitation à diriger des recherches nécessite un contrôle cognitif.

La notion de contrôle cognitif est l'objet de suffisamment de confusions pour justifier que l'on s'arrête quelques instants sur la terminologie. D'autres appellations s'utilisent de façon plus ou moins interchangeable, et renvoient moins à des divergences conceptuelles qu'à des différences d'usage et de traditions de recherche : par exemple, *contrôle exécutif* ou *fonctions exécutives* relèvent plutôt de la tradition neuropsychologique et suggèrent une multiplicité de fonctions exécutives réalistes qui sont autant de formes de contrôle différentes. Le terme de *contrôle attentionnel* relève plutôt de la recherche sur la mémoire de travail et implique un rôle fonctionnel central de l'attention dans la régulation (qui ne se justifie pas forcément au regard des travaux montrant que le contrôle est possible en l'absence d'attention ou même de conscience du participant ; Hommel, 2007, 2013 ; Van Gaal et al., 2012). Je privilégie ici le terme *contrôle cognitif* qui semble, par contraste, relativement agnostique quant à la nature et aux déterminants du contrôle mis en place.



"Réguler son activité cognitive dans une situation qui ne peut pas être résolue de façon automatique" : cette définition a de quoi alarmer le psychologue différentialiste, pour deux raisons symétriques. D'une part, c'est une définition assez vaste pour conférer au contrôle cognitif un rôle potentiel dans l'ensemble des activités cognitives de haut niveau. Autrement dit, l'aptitude au contrôle cognitif est susceptible d'être source de variabilité dans à peu près n'importe quelle tâche complexe. La littérature n'a d'ailleurs pas manqué de mettre en lumière des liens entre contrôle cognitif et performance, dans les situations les plus diverses (Diamond, 2013, donne par exemple une liste qui va de la satisfaction maritale à la réussite à l'école en

passant par l'obésité). D'autre part, il est question de la régulation, du contrôle d'autres fonctions cognitives : on s'attend donc à ce que la variabilité dans une situation qui nécessite un contrôle cognitif soit intrinsèquement co-déterminée non seulement par l'aptitude au contrôle cognitif, mais aussi par les fonctions cognitives qui font l'objet d'un contrôle. Autrement dit, la mesure du contrôle cognitif est par définition indissociable de la mesure d'autres fonctions cognitives, l'efficacité du contrôle cognitif n'étant apparente qu'à travers la bonne exécution de ces autres fonctions.



Si le contrôle cognitif est indissociable des processus cognitifs qu'il permet de réguler, comment être sûr qu'il s'agit d'un construit à part entière, qu'il serait justifié d'étudier pour lui-même ? La recherche s'est beaucoup appuyée sur l'existence d'atteintes cliniques spécifiques au contrôle cognitif, prenant la forme d'un syndrome dysexécutif ("syndrome frontal") : les patients affectés, habituellement à la suite d'une lésion du cortex préfrontal, présentent un trouble général de la régulation comportementale et cognitive, associant par exemple une difficulté à inhiber les comportements ou émotions inadaptées, des persévérations, ou encore un défaut de planification. Le cas clinique le plus connu est celui de Phineas Gage ; même si ses symptômes semblent avoir été largement exagérés par le folklore scientifique (Macmillan 2008), l'existence de ce type de troubles exécutifs peut être vue comme la validation directe de la réalité phénoménologique d'une fonction de contrôle cognitif.

Face à ces atteintes globales du contrôle cognitif, les premiers travaux de recherche ont généralement abordé le contrôle cognitif comme une aptitude unitaire, dont tous les effets seraient sous-tendus par une unique entité réaliste localisée au niveau du cortex préfrontal. Le modèle du superviseur attentionnel de Norman et Shallice (1986) ou encore l'administrateur central de Baddeley (1996) sont des exemples de cette conception. Si ces modèles ont permis de fournir un premier cadre interprétatif et d'accompagner les premières recherches sur le contrôle cognitif, ils souffrent d'un défaut rédhibitoire : ils ne permettent pas de spécifier le fonctionnement du contrôle mis en jeu. Le contrôle cognitif est conçu comme étant implémenté par un "administrateur central" qui "administre" le fonctionnement cognitif ; une atteinte exécutive correspond à un dysfonctionnement de "l'administrateur central" ; à l'inverse, une bonne performance exécutive reflète un bon fonctionnement de "l'administrateur central". Autrement dit, ces modèles ont un intérêt descriptif, mais non explicatif : le contrôle cognitif prend la forme d'un *homoncule* – un petit bonhomme à

l'intérieur de l'esprit qui s'occupe des opérations de contrôle et dont le fonctionnement n'est pas ou peu spécifié (e.g. Logie, 2016). En l'état, ce type de conception n'est pas falsifiable et présente donc un intérêt heuristique assez limité pour qui cherche à mieux comprendre la variabilité du contrôle.



Une seconde série de travaux a cherché à mieux comprendre la structure du contrôle cognitif en listant les différentes fonctions qu'il assure, et en distinguant des tâches destinées à mesurer l'efficacité de chacune de ces fonctions exécutives. Les travaux les plus connus dans ce domaine sont ceux de Miyake et collaborateurs (2000), dont la contribution majeure a été de tester empiriquement l'existence de diverses fonctions exécutives présumées distinctes en faisant passer un ensemble de 14 tâches supposées évaluer différentes fonctions, puis en cherchant des regroupements à l'aide d'analyses factorielles. Cette approche a conduit les auteurs à distinguer trois fonctions exécutives majeures : inhibition, flexibilité mentale, et mise à jour en mémoire. À travers l'évaluation d'aptitudes distinctes, ce modèle permet un pas en avant dans l'évaluation de la variabilité, et peut être considéré comme la référence contemporaine concernant l'évaluation du contrôle cognitif parmi les non-spécialistes, en particulier dans la littérature neuropsychologique. La démarche consistant à distinguer un ensemble de fonctions exécutives pose toutefois un grand nombre de problèmes conceptuels et pratiques qui justifient, à mon sens, un certain scepticisme quant à son utilisation dans une perspective différentielle (Gonthier, 2017).

Sur le plan conceptuel, on relèvera la grande variabilité de la liste des fonctions exécutives identifiées (souvent de façon purement subjective) d'un auteur à l'autre : *formuler des buts, planifier des actions, les implémenter, et le faire efficacement* (Lezak, 1982) ; mais aussi *réguler ses émotions, sa motivation et son degré d'éveil ; analyser et synthétiser une situation ; maintenir des informations en mémoire de travail* (Barkley, 1997) ; ou encore *la planification ; l'attention sélective ; la fluence verbale et comportementale ; la résolution de problème* (Anderson et al., 2008) ; voire *l'aptitude à réaliser deux tâches simultanément* (Miyake et al., 2000). La contribution majeure des travaux de Miyake et collaborateurs (2000) a été d'éviter le caractère subjectif de ces listes à la Prévert grâce à une étude empirique des regroupements factoriels ; mais cette approche se heurte à une limite bien connue de l'analyse factorielle (e.g. Thurstone, 1940) : les facteurs que l'on extrait sont intimement dépendants des tâches que l'on a choisi d'inclure dans l'analyse (une limite soulevée par les auteurs eux-

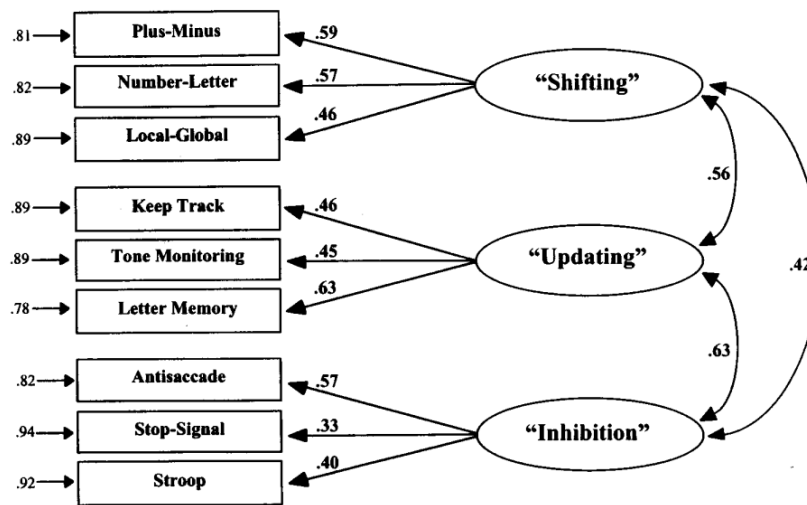
mêmes)... limite doublement problématique au vu de la grande diversité des tâches exécutives. Ainsi, des choix de tâches différents auraient conduit à identifier des fonctions différentes.

Mais le problème conceptuel de loin le plus grave a trait au fait que la définition des fonctions exécutives dépend exclusivement des pré-conceptions des chercheurs quant à la nature des processus mis en jeu par la tâche. Ainsi, la fonction exécutive d'*inhibition*, probablement la plus universellement mentionnée dans la littérature, n'a pas de fondement formel évident : quoique cette fonction ait largement été justifiée par l'existence de processus inhibiteurs au niveau neurologique, sur le plan cognitif, la totalité des résultats attribués à une fonction d'inhibition peuvent être réinterprétés en termes d'activation différenciée de deux processus, le contrôle cognitif consistant à renforcer l'activation de l'un et non à inhiber l'activation de l'autre (MacLeod et al., 2003). Des problèmes similaires se posent pour les autres fonctions exécutives. Ce problème est renforcé par les chevauchements conceptuels entre fonctions exécutives : ainsi, la fonction de mise à jour peut être vue comme nécessitant l'inhibition de la représentation mentale précédente pour pouvoir accueillir la suivante, tout comme la flexibilité mentale peut être vue comme une forme de mise à jour... Il semble donc peu pertinent de tracer de distinctions absolues entre fonctions comme si elles étaient parfaitement distinctes. L'un dans l'autre, les fonctions identifiées par la littérature apparaissent donc plutôt comme un cadre interprétatif inventé par les psychologues pour catégoriser les tâches qu'ils utilisent, sans nécessairement recouvrir de réalité ontologique pour le participant : on atteint ici la limite de la conception réaliste des fonctions exécutives.



Sur le plan pratique, l'étude de fonctions exécutives distinctes révèle une grande instabilité des résultats d'une étude à l'autre, les regroupements factoriels variant au gré des échantillons, des tâches et du hasard pour faire apparaître deux fonctions exécutives (e.g. Hedden & Yoon, 2006 ; éventuellement assorties d'une structure hiérarchique ; Miyake & Friedman, 2012), ou une seule (e.g. Wiebe et al., 2008), ou quatre (Bock et al., 2019). En plus des problèmes conceptuels évoqués précédemment, cette instabilité s'explique en partie par les faibles qualités psychométriques des tâches exécutives (Miyake et al., 2000 ; Rabbit, 1997), dont la performance se calcule souvent comme un score de différence dont la fidélité tend à être limitée (Thomas & Zumbo, 2012) ; elle s'explique aussi par les faibles corrélations entre tâches (Bock et al., 2019). Ainsi, si l'analyse factorielle confirmatoire de Miyake et collaborateurs est interprétée comme validant l'existence de trois fonctions exécutives

(Figure 1a), l'examen des corrélations bivariées reléguées en annexe par les auteurs révèle que des tâches supposées mesurer la même fonction exécutive sont en réalité très faiblement liées entre elles (Figure 1b) : par exemple, les trois tâches mesurant "l'inhibition" (antisaccade, stop-signal, Stroop) corréleront à .19, .20 et .18. Au-delà de la faible fidélité des tâches exécutives, ces faibles corrélations entre tâches s'expliquent peut-être, en partie, par le caractère intrinsèquement impur des mesures (Rabbitt, 1997), problème discuté un peu plus haut : ainsi, un test de Stroop mesure non seulement le contrôle cognitif ou l'inhibition, mais aussi la vitesse de lecture et la vitesse de dénomination des couleurs ; il est donc logique de trouver une importante variance tâche-spécifique.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Plus-minus	—														
2. Number-letter	.32*	—													
3. Local-global	.23*	.32*	—												
4. Keep track	.23*	.08	.12	—											
5. Tone monitoring	.22*	.19*	.00	.15	—										
6. Letter memory	.24*	.11	.21*	.34*	.27*	—									
7. Antisaccade	.15	.17	.11	.12	.26*	.22*	—								
8. Stop-signal	.11	.13	.06	.10	.09	.04	.19*	—							
9. Stroop	.07	.09	-.05	.11	.16	.18*	.20*	.18*	—						
10. WCST perseveration ^a	.26*	.13	.18*	.09	.19*	.14	.15	-.01	.10	—					
11. TOH ^b	.08	.10	-.09	.13	.18*	.14	.21*	.08	.17	-.02	—				
12. RNG Component 1	.20*	.13	.01	.03	.11	.19*	.24*	.12	.11	.13	.10	—			
13. RNG Component 2	.20*	-.07	.07	.29*	.06	.19*	.02	.18*	.01	-.08	.12	.02	—		
14. Operation span	.09	.08	-.04	.41*	.28*	.34*	.16	.13	.20*	.16	.04	.17*	.13	—	
15. Dual task ^a	-.03	-.02	.05	-.09	-.03	.12	-.08	-.16	.06	.06	-.18*	-.05	-.09	-.14	—

Note. WCST, Wisconsin Card Sorting Test; TOH, Tower of Hanoi; RNG, random number generation.

^a N = 134.

^b N = 136.

* p < .05.

Figure 1. Structure des fonctions exécutives à travers les résultats d'une analyse factorielle confirmatoire et à travers les corrélations bivariées correspondantes. Tiré de Miyake, A., et al. (2000), The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, 41, 49-100.

Enfin, distinguer une multiplicité de fonctions exécutives pose un problème pratique au scientifique : mesurer trois fonctions exécutives, c'est s'exposer trois fois à un risque de faux positif, et c'est aussi risquer de devoir interpréter des résultats divergeant d'une fonction à l'autre sans explication théorique apparente.

La démarche consistant à distinguer des fonctions exécutives peut être vue comme une approche structurelle du contrôle cognitif : une tentative de spécifier la structure du construit en délimitant des boîtes de rangement distinctes dans lesquelles on peut classer ses différentes manifestations ou composantes. Une alternative possible est d'adopter une démarche fonctionnelle : s'intéresser aux mécanismes qui permettent d'implémenter le contrôle cognitif en pratique (Gonthier, 2017). L'accent est alors mis sur le fonctionnement du processus de contrôle : que signifie "contrôler" son activité cognitive, selon quelles modalités ce contrôle est-il mis en jeu, à quel moment, et avec quels effets ? Dans ce contexte, le passage d'une démarche structurelle à une approche fonctionnelle tend à aller de pair avec le passage d'une approche corrélationnelle (visant à rechercher des regroupements factoriels entre tâches, de façon à spécifier une structure interne du contrôle) à une approche expérimentale (visant à manipuler expérimentalement les modalités d'implémentation du contrôle de façon à mieux comprendre ses contraintes et ses effets).



Le modèle de Miller et Cohen (2001) est à mon avis la plus incontournable des conceptions fonctionnelles du contrôle cognitif. Cette conception présente le contrôle cognitif comme un processus consistant à utiliser la représentation du but de la tâche, stockée par le cortex préfrontal, pour renforcer sélectivement une association entre stimulus et réponse, de façon à biaiser le traitement cognitif en faveur du comportement attendu. Une application de ce modèle connexionniste à la tâche de Stroop est représentée dans la Figure 2. Le modèle propose que les dimensions perceptives d'un stimulus tel que **vert** (le mot *vert*, la couleur ■) sont liées à des réponses possibles ("vert", "rouge"). Un traitement automatique de cette tâche aboutirait à donner la réponse prédominante, mais incorrecte, et donc à lire le mot : "vert". Appliquer un contrôle cognitif, c'est maintenir activement en mémoire le but de la tâche ("je dois dénommer la couleur") par l'intermédiaire du cortex préfrontal, de façon à renforcer sélectivement l'association entre la dimension pertinente (la couleur ■) et la réponse correspondante ("rouge"). Renforcer cette association conduit à biaiser le traitement en faveur de la réponse attendue mais initialement non dominante, et donne lieu au contrôle cognitif.

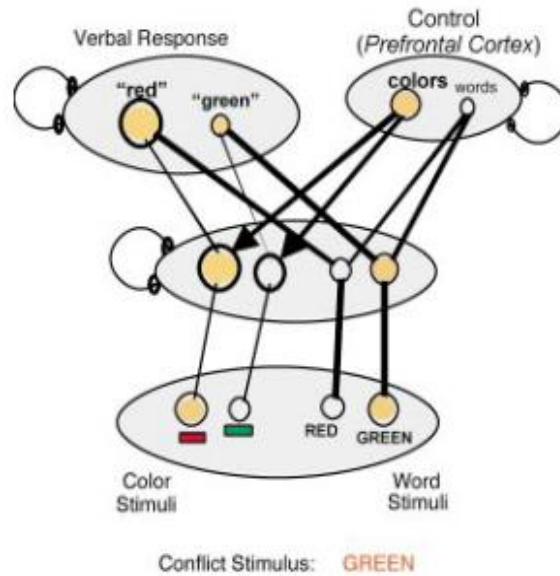


Figure 2. Modèle fonctionnel connexionniste du contrôle cognitif dans la tâche de Stroop. Tiré de Miller, E. K., et Cohen, J. D. (2001), An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24, 167-202.

Ce type de modèle fonctionnel autorise plusieurs façons de s'intéresser à la variabilité des performances de contrôle. Les travaux de Todd Braver et collaborateurs (Braver et al., 2007) ont proposé un angle d'attaque original sous la forme du modèle *Dual Mechanisms of Control* (DMC). Le modèle DMC propose l'existence de deux mécanismes possibles d'implémentation du contrôle cognitif : contrôle proactif, et contrôle réactif. Le contrôle proactif consiste à maintenir activement le but de la tâche en mémoire de travail avant même l'apparition du stimulus qui doit faire l'objet du contrôle, de façon à pouvoir engager le contrôle cognitif rapidement lorsque le stimulus en question apparaît. Cette forme de contrôle repose sur une activation soutenue du cortex dorsolatéral préfrontal, dont le rôle consiste à maintenir activement le but de la tâche. À l'inverse, le contrôle réactif consiste à réactiver sélectivement le but de la tâche lorsqu'un stimulus nécessitant un contrôle apparaît, ce qui aboutit à mettre en jeu une régulation plus tardive. Cette forme de contrôle donne plutôt un rôle central au cortex cingulaire antérieur, qui joue un rôle de détection des situations de conflit nécessitant un contrôle cognitif ; la présence d'une situation de conflit conduit à réactiver le cortex dorsolatéral préfrontal pour implémenter le contrôle.



Cette différence de temporalité entre contrôle proactif et contrôle réactif est à l'origine d'un ensemble d'avantages et d'inconvénients pour ces deux mécanismes, qui autorisent des



prédictions fonctionnelles assez détaillées. Le contrôle proactif est vu comme étant plus efficace dans la plupart des situations, puisqu'étant mis en jeu à l'avance ; en contrepartie, cette forme de contrôle est supposée plus coûteuse en "ressources" (cognitives ou métaboliques), dans la mesure où elle demande une activation soutenue du cortex dorsolatéral préfrontal pour maintenir activement le but de la tâche. Par ailleurs, ce mécanisme tend à être plus rigide puisque s'appliquant de façon globale (sans discrimination) sur une durée étendue, et il ne peut être efficace que si la situation intègre des indices qui permettent de prédire de façon fiable le comportement qui devra être mis en place lorsque le stimulus critique apparaîtra. À l'inverse, le contrôle réactif tend à être moins efficace puisque mis en jeu plus tardivement, et il est plus vulnérable lorsque le stimulus présenté tend à induire une réponse incorrecte. Le contrôle réactif est en revanche plus flexible puisque mis en jeu uniquement quand un stimulus le nécessite, et il est aussi moins dépendant de la présence d'indices prédictifs fiables.

Le modèle DMC propose que la variabilité des performances de contrôle cognitif repose largement sur une variation qualitative dans la sélection de l'un ou l'autre mode de contrôle. Ainsi, la performance élevée de contrôle cognitif chez le jeune adulte sain serait due à l'usage préférentiel du contrôle proactif, tandis que la plus faible performance de contrôle cognitif observée dans le cadre du vieillissement normal ou pathologique ou encore de la schizophrénie serait due à une transition progressive vers le contrôle réactif. Cette transition est généralement attribuée à un déclin de l'intégrité du substrat neuronal nécessaire à la maintenance active du but de la tâche, requise pour déployer un contrôle proactif (e.g. Braver et al., 2005 ; Paxton et al., 2006).

Pour le différentialiste francophone, cette conception qualitative de la variabilité comme reposant sur la sélection d'un processus parmi plusieurs options renvoie bien sûr à la notion de processus vicariants défendue par Maurice Reuchlin (1978) ; mais ce n'est pas le seul attrait de ce modèle. Les patterns d'effets positifs et négatifs proposés par les auteurs constituent autant de signatures comportementales des deux mécanismes de contrôle. Ces hypothèses fonctionnelles ouvrent ainsi la voie à une identification précise du mécanisme mis en jeu par un participant particulier dans une tâche particulière, et donc à l'identification d'un déterminant causal de la variabilité qui serait valable au niveau individuel.

Le modèle DMC propose ainsi une plus grande finesse dans l'analyse des causes de la variabilité que celle traditionnellement offerte par les approches en fonctions exécutives, qui se limitent généralement à une lecture quantitative simple ("bonne inhibition", "mauvaise inhibition"). Ce modèle n'est pas exempt de problèmes (dont une certaine tendance des auteurs à la spéculation quant au fonctionnement des deux mécanismes, sur une base logique mais pas forcément empirique), mais présente donc un intérêt heuristique certain pour l'analyse des déterminants de la performance dans les tâches exécutives. Il a, de fait, été utilisé dans un grand nombre d'études visant à comprendre les différences entre groupes, mais aussi la variation inter-individuelle ou intra-individuelle du contrôle (Braver et al., 2007 ; Braver, 2012).

Dans le cadre de mes travaux de doctorat (Gonthier, 2014), j'ai par exemple cherché à m'appuyer sur le modèle DMC pour interpréter l'origine de la performance de contrôle cognitif plus élevée des participants qui présentent une bonne mémoire de travail, avec l'hypothèse que ces participants utilisent préférentiellement le contrôle proactif. Si mon travail doctoral a radicalement échoué à vérifier cette hypothèse (d'autres auteurs ayant pourtant, depuis, argumenté en ce sens : e.g. Redick & Engle, 2011 ; Redick, 2014 ; Richmond et al., 2015), il m'a en tout cas convaincu de l'intérêt de ce modèle pour l'interprétation de la variabilité.

2. Évaluer les mécanismes de contrôle

2.1. Le modèle DMC à travers l'AX-CPT

Pour utiliser le modèle DMC dans l'étude de la variabilité, un enjeu majeur est de pouvoir évaluer quel mécanisme de contrôle – proactif ou réactif – est utilisé par un participant donné. Le développement du modèle s'est presque exclusivement appuyé sur une tâche en particulier, l'AX-CPT. Cette tâche est dérivée du paradigme de Continuous Performance Test (CPT ; Rosvold et al., 1956), destiné à évaluer l'attention soutenue. La version originelle du CPT présentait une suite de lettres, le participant ayant pour consigne d'appuyer sur la touche réponse à chaque fois qu'il voyait apparaître la lettre X. Le principe du paradigme AX-CPT est plus complexe et introduit une dimension temporelle : le participant a pour consigne de donner une réponse "oui" à chaque fois qu'il voit apparaître la lettre X, mais seulement si la lettre précédente était un A. Cette modification rend possible différentes façons de résoudre la tâche, en anticipant ou non sa réponse à la lettre suivante.

Les essais sont habituellement découpés en séries de deux lettres, dont la présentation est séparée par un court délai. La première lettre est nommée l'*indice* et la seconde la *cible*. On distingue classiquement quatre types d'essais : AX (la lettre A suivie de la lettre X), AY (la lettre A suivie d'une autre cible que X), BX (un autre indice que A suivi d'un X), et BY (un autre indice que A suivi d'une autre cible que X). La clef de voûte du paradigme est le fait que les essais AX sont particulièrement fréquents (70% de tous les essais dans la version la plus courante, contre 10% pour chacun des trois autres types d'essais) ; autrement dit, la probabilité conditionnelle de voir apparaître la cible X lorsque l'indice était un A est particulièrement élevée. Il est donc possible d'engager un contrôle proactif : lorsque l'indice est un A, on peut préparer une réponse "oui" qui a de fortes chances d'être correcte, et maintenir activement la préparation de cette réponse pendant le délai qui précède l'apparition de la cible ; à l'inverse, lorsque l'indice n'est pas un A, on peut se préparer à donner une réponse "non" qui sera nécessairement correcte. Le contrôle réactif consiste simplement à attendre l'apparition de la cible pour sélectionner sa réponse, en récupérant si besoin l'identité de l'indice en mémoire.

L'avantage majeur de l'AX-CPT, qui en fait aussi une tâche assez unique en psychologie cognitive, est le caractère différencié des prédictions que ce paradigme autorise en fonction du mécanisme de contrôle utilisé. Le contrôle proactif doit amener à des réponses plus rapides et

plus souvent correctes pour les essais AX, mais surtout pour les essais BX : l'indice B permet en effet de préparer à l'avance la réponse "non" et de la donner presque instantanément lorsque la cible est présentée, sans risque d'être induit en erreur par le fait qu'il s'agit d'un X. En revanche, le contrôle proactif doit amener à des réponses plus lentes et à un risque d'erreur important sur les essais AY, l'indice A conduisant à préparer la réponse "oui" qui est la plus probable mais s'avère erronée lorsque la cible Y apparaît. L'usage d'un contrôle proactif doit donc se traduire par une performance élevée en BX et faible en AY, tandis que le pattern inverse s'observe pour le contrôle réactif. (Un exemple de patterns proactif et réactif est présenté nettement plus loin dans ce manuscrit, Figure 5, page 49 ; preuve à la fois d'un sens de l'économie dans les figures et d'une construction logique vacillante de ce travail).



L'AX-CPT présente ainsi une rare situation dans laquelle on peut observer une performance *plus faible* pour les participants qui utilisent le mécanisme de contrôle *le plus efficace*. C'est un avantage précieux pour l'évaluation des causes de la variabilité, qui permet de conclure avec une relative certitude quant à l'origine des différences observées : si le taux d'erreurs d'un jeune adulte sur les essais AY est plus élevé que celui d'une personne âgée atteinte de la maladie d'Alzheimer (Braver et al., 2005), on peut croire avec les auteurs que l'origine de cette différence est bien l'usage du contrôle proactif dans un cas et réactif dans l'autre. Par comparaison, la quasi-totalité des tâches classiques de contrôle cognitif (tâche de Stroop, go/no-go, tâche d'antisaccades, task-switching, *n*-back, et ainsi de suite) révèlent uniquement que les participants les plus efficaces ont une meilleure performance, ce qui ne permet pas de conclusions fines quant aux mécanismes à l'origine de cette différence.



L'AX-CPT autorise le calcul de plusieurs indices de performance. Taux d'erreur et temps de réponse sur les quatre types d'essais sont bien sûr au premier plan de l'analyse, mais l'interprétation repose généralement sur un indice composite censé refléter le degré d'usage du contrôle proactif. La littérature en propose trois, dont deux s'appuient sur la théorie de la détection du signal (TDS). Le *d'*-context est inspiré de l'indice de sensibilité *d'* en TDS, et se calcule sous la forme $z(\text{bonnes réponses AX}) - z(\text{erreurs BX})$ [où *z* désigne la fonction inverse de la fonction de distribution de la loi normale]. Cet indice reflète l'efficacité dans l'utilisation de l'indice contextuel pour déterminer la réponse faite à une cible X. Quelques travaux ont également utilisé le *A-cue bias*, inspiré de l'indice de biais *c* en TDS, et se calcule sous la forme $0.5 * [z(\text{bonnes réponses AX}) + z(\text{erreurs AY})]$. Cet indice estime la tendance à préparer une

réponse "oui" lorsque l'indice est un A. Enfin, le *proactive behavioral index* ou *PBI* se calcule sous la forme $(AY - BX) / (AY + BX)$. Cet indice peut se calculer pour les taux d'erreur ou pour les temps de réponse ; on peut d'ailleurs retenir ces deux solutions et en faire la moyenne, après standardisation².

Des trois indices composites censés refléter l'usage du contrôle proactif, le PBI est à mon sens le plus intéressant, ce qui m'a amené à lui donner une place prépondérante dans mes travaux utilisant l'AX-CPT (e.g. Gonthier, Macnamara, Chow, Conway, & Braver, 2016; Gonthier, Zira, Colé, & Blaye, 2019), pour deux raisons majeures. 1) Comme sa formule le montre, le PBI reflète directement l'équilibre entre performance sur les essais AY et BX, qui est au cœur des prédictions du modèle DMC ; par comparaison, le *d'*-context beaucoup plus fréquent dans la littérature ne tient pas compte des essais AY et permet donc de conclure fallacieusement d'un participant globalement plus efficace qu'il "utilise le contrôle proactif". 2) Contrairement aux autres indices, le PBI peut être calculé aussi bien pour les temps de réponse que pour les taux d'erreur, voire pour la combinaison des deux, ce qui permet de contrôler la possibilité de tradeoffs entre vitesse et précision ; là encore, un indice comme le *d'*-context n'examinant que les taux d'erreur amènera à conclure d'un participant plus précautionneux qu'il "utilise le contrôle proactif".



L'existence d'un si grand nombre d'indices de performance pour une même tâche est en soi problématique, à la fois d'un point de vue conceptuel (les trois indices composites sont considérés comme interchangeables par la littérature, en dépit de raisonnements très différents ; on pourrait argumenter, par exemple, que le *d'*-context représente l'efficacité du contrôle plutôt que la nature du mécanisme de contrôle utilisé) et d'un point de vue pratique (trois indices de performance, ce sont autant de "degrés de liberté chercheur" qui permettent le cherry-picking de l'indice qui fonctionne le mieux). Mais ce n'est pas le seul problème de l'AX-CPT. Une autre limite majeure de ce paradigme pour l'étude de la variabilité réside dans ses faibles qualités psychométriques.

À l'occasion d'une série de quatre études évaluant systématiquement les propriétés métriques de l'AX-CPT (Cooper, Gonthier, Barch, & Braver, 2017), nous avons montré que la tâche présente habituellement des coefficients de fidélité très limités. Ainsi, nous avons pu

² On peut relever que ces indices composites se calculent après application d'une correction log-linéaire, qui les rend calculables même lorsque les taux d'erreur sont de 0 ou de 1 ; voir Gonthier et al., 2016.

relever des coefficients alpha de Cronbach ou des corrélations test-retest proches de zéro pour des versions pourtant classiques du paradigme, et ce qu'il soit réalisé en laboratoire ou à distance. Ce problème se pose aussi bien pour les indices composites, que pour les mesures plus simples des taux d'erreur. Si cette faible fidélité n'était pas forcément rédhibitoire pour les premiers travaux utilisant l'AX-CPT et qui consistaient en de simples comparaisons inter-groupes (selon la logique consistant à affirmer qu'une faible fidélité équivaut à une variance d'erreur plus importante, donc à une moindre puissance statistique pour réaliser une comparaison de moyennes), elle constitue évidemment une limite majeure pour une étude de la variabilité inter ou intra-individuelle qui, par nature, nécessite une mesure précise de la position de chaque individu.

Une analyse détaillée de la structure de la tâche et des résultats obtenus nous a permis de situer les principales causes de cette faible fidélité (Cooper, Gonthier, Barch, & Braver, 2017). D'une part, la tâche tend à présenter un effet plafond pour les populations adultes saines, pour lesquelles le contrôle proactif est le mécanisme utilisé par défaut. La plupart des participants commettent peu ou pas d'erreurs, ce qui donne lieu à des mesures à faible variance, qui limitent mécaniquement les coefficients de fidélité par un phénomène de restriction d'étendue (la sensibilité constituant un prérequis de la fidélité). D'autre part, les deux types d'essais critiques sont les essais AY et BX ; mais ces deux essais sont aussi les moins fréquents (10% du total pour chacun). Ainsi, les versions courantes du paradigme, qui intègrent une centaine d'essais, n'offrent en fait que 10 essais permettant d'évaluer la performance en AY ou BX. Ce faible nombre d'essais limite naturellement la précision de la mesure, mais pose un problème supplémentaire pour les taux d'erreurs : le nombre de scores différents possibles est assez limité, ce qui amplifie encore le problème de sensibilité causé par la performance élevée des participants (dont les taux d'erreurs observés se limitent habituellement à 0.00, 0.10 ou 0.20). Les indices composites comme le PBI posent un problème supplémentaire : tous se calculent comme des scores de différences (AX vs. BX, AX vs. AY ou encore AY vs. BX), dont on sait qu'ils tendent à présenter une fidélité faible, la variance d'erreur d'une différence entre scores étant fonction de la variance d'erreur de chacun des deux scores considérés séparément (Thomas & Zumbo, 2012).

En pratique, nos analyses ont montré qu'environ 200 essais étaient nécessaires pour approcher un seuil de fidélité acceptable pour les essais BX et BY qui donnent lieu à peu

d'erreurs chez les adultes sains ; mais cette conclusion n'est pas très encourageante, car il est difficile de convaincre les participants de réaliser autant d'essais d'une tâche relativement ennuyeuse (le délai entre cible et indice étant généralement autour de 5 secondes, on comprend que le temps puisse leur paraître long). La plupart des recherches utilisant l'AX-CPT s'appuient donc sur un nombre d'essais limité (tout en discutant rarement les propriétés psychométriques de leur mesure, qui bien sûr dépendent de l'échantillon), ce qui limite le crédit que l'on peut donner aux résultats corrélacionnels et contribue probablement aux difficultés de réplification rapportées par la littérature. Nos travaux ont d'ailleurs permis de montrer (Cooper, Gonthier, Barch, & Braver, 2017) que l'instabilité de certains résultats d'une étude à l'autre étaient dus en partie à des différences dans la composition des échantillons, qui génèrent des différences de sensibilité, donc de fidélité, et donc imposent des limites différentes sur la taille maximale des corrélacions qui peuvent être observées. Deux bons exemples sont l'étude des corrélacions entre usage du contrôle proactif et mémoire de travail chez l'adulte sain, décrites par la littérature (Redick & Engle, 2011 ; Redick, 2014 ; Richmond et al., 2015) mais que nous avons échoué à répliquer ; et l'étude des corrélacions entre usage du contrôle proactif et mémoire épisodique, observées dans un échantillon de patients schizophrènes mais non dans un échantillon de participants sains. Dans les deux cas, l'échantillon dans lequel les corrélacions attendues n'étaient pas observées présentait un niveau élevé de contrôle proactif, donc des mesures à la sensibilité faible, et des coefficients de fidélité également faibles.

Ce problème de mesure n'est pas spécifique à l'AX-CPT et concerne de nombreuses tâches exécutives qui présentent un fonctionnement similaire (essais rares, comme dans le go/no-go, ou calcul de scores de différence, comme dans le test de Stroop ou le Trail Making Test). À cette limite psychométrique s'ajoute toutefois une considération plus théorique : la quasi-totalité de la littérature concernant le modèle DMC a utilisé diverses variantes de l'AX-CPT, ce qui limite la généralisabilité du modèle. L'AX-CPT fait en effet appel à plusieurs mécanismes (lecture de lettres, rapidité motrice...) qui lui sont spécifiques, et qui peuvent constituer autant de variables confondues contribuant aux résultats obtenus. Ce point nous a encouragés à rechercher d'autres situations permettant d'évaluer le contrôle proactif et le contrôle réactif. J'ai essentiellement travaillé sur les effets de proportion de congruence.



3.2. Le modèle DMC à travers les effets de proportion de congruence

Les tâches de conflit confrontent le participant à des stimuli qui présentent deux dimensions perceptives : une dimension qui doit déterminer la réponse, et une dimension qui doit être ignorée. Dans le test de Stroop, les participants doivent identifier la couleur dans laquelle est écrit un nom de couleur, sans prêter attention au mot lui-même ; dans la flanker task de type flèches, les participants doivent identifier la direction d'une flèche présentée au centre de l'écran (par exemple >), sans prêter attention à d'autres flèches présentées de part et d'autre et appelées flankers (par exemple < < > < <). Ces deux dimensions perceptives peuvent être compatibles ou incompatibles ; lorsqu'elles sont incompatibles, on observe un conflit, une interférence, qui peut donner lieu à un ralentissement de la réponse et à des erreurs. On considère que le contrôle cognitif est requis dans les essais incongruents pour favoriser le traitement de la dimension perceptive à laquelle la tâche demande de répondre, tout en évitant de traiter la dimension non pertinente. Ces tâches de conflit peuvent donner lieu à des effets de proportion de congruence, qui offrent une fenêtre possible sur le contrôle proactif et le contrôle réactif.

Le principe général des effets de proportion de congruence est le suivant : lorsque la plupart des essais présentés au sein de la tâche sont incongruents, on observe une interférence moindre que lorsque la plupart des essais sont congruents. Tout se passe comme si, étant régulièrement confrontés à des stimuli dont les deux dimensions sont incompatibles, les participants apprenaient à appliquer un contrôle cognitif permettant de réduire le poids accordé à la dimension non pertinente ou d'augmenter le poids accordé à la dimension pertinente, de façon à réduire l'interférence ressentie. L'intérêt des effets de proportion de congruence dans le cadre du modèle DMC est qu'il existe plusieurs façons de les manipuler, et dont nous avons argumenté qu'elles reflètent soit le contrôle proactif, soit le contrôle réactif (Gonthier, Braver, & Bugg, 2016). On peut distinguer trois types d'effets de proportion de congruence, résumés dans la Figure 3 :

1) L'effet List-Wide Proportion Congruency (LWPC) s'observe en comparant un bloc d'essais dans lequel la plupart des essais sont incongruents (Mostly Incongruent, MI) à un bloc d'essais dans lequel la plupart des essais sont congruents (Mostly Congruent, MC) : l'interférence est réduite dans le bloc MI par rapport au bloc MC. Un point critique est que cette réduction de l'interférence s'observe même pour des stimuli qui sont congruents 50% du temps au sein des

deux blocs d'essais, ce qui indique que la réduction de l'interférence relève d'un mécanisme de contrôle cognitif appliqué de façon globale, à travers tous les essais du bloc. Par définition, il s'agit donc d'une forme de contrôle soutenue dans le temps, et appliquée avant même l'apparition d'un stimulus à travers tout un bloc d'essais ; on peut donc la qualifier de proactive (e.g. Hutchison, 2011).

2) L'effet Item-Specific Proportion Congruency (ISPC) s'observe en comparant des stimuli généralement incongruents (MI) à des stimuli généralement congruents (MC). Les stimuli MI et MC sont définis sur la base de la dimension pertinente, à laquelle le participant doit répondre : par exemple, dans une tâche de Stroop, sur la base de la couleur du mot³. Ainsi, les mots écrits en rouge et en bleu pourraient être MI et les mots écrits en vert et en jaune pourraient être MC. On observe que l'interférence est réduite pour les stimuli MI par rapport aux stimuli MC. Les stimuli MI et MC étant mélangés au sein d'un même bloc d'essai, cette réduction de l'interférence ne peut être due qu'à un mécanisme appliqué de façon flexible après l'apparition du stimulus nécessitant un contrôle. Par définition, il s'agit donc d'une forme de contrôle implémentée après coup, que l'on peut qualifier de réactive (e.g. Hutchison, 2011).

3) L'effet de Context-Specific Proportion Congruency (CSPC) est un peu à part et s'observe également en comparant des stimuli MI et MC, la différence avec l'effet ISPC étant que les sets de stimuli MI et MC sont définis sur la base d'un aspect du stimulus qui n'a aucun lien direct avec la tâche et ne doit évoquer aucune réponse particulière : par exemple, dans une tâche de Stroop, sur la base de la police de caractères (Bugg et al., 2008). L'effet CSPC a été nettement moins étudié mais, tout comme l'effet ISPC, il ne peut logiquement s'attribuer qu'à un mécanisme de contrôle réactif.

³ Il est également possible de définir les stimuli MI et MC sur la base de la dimension non pertinente, mais cela génère des explications alternatives en termes d'apprentissage associatif de contingences, l'interférence pouvant être réduite pour les stimuli MC parce que les sujets apprennent la réponse la plus fréquente à donner lorsqu'ils lisent un mot généralement congruent ; voir Schmidt & Besner, 2008.

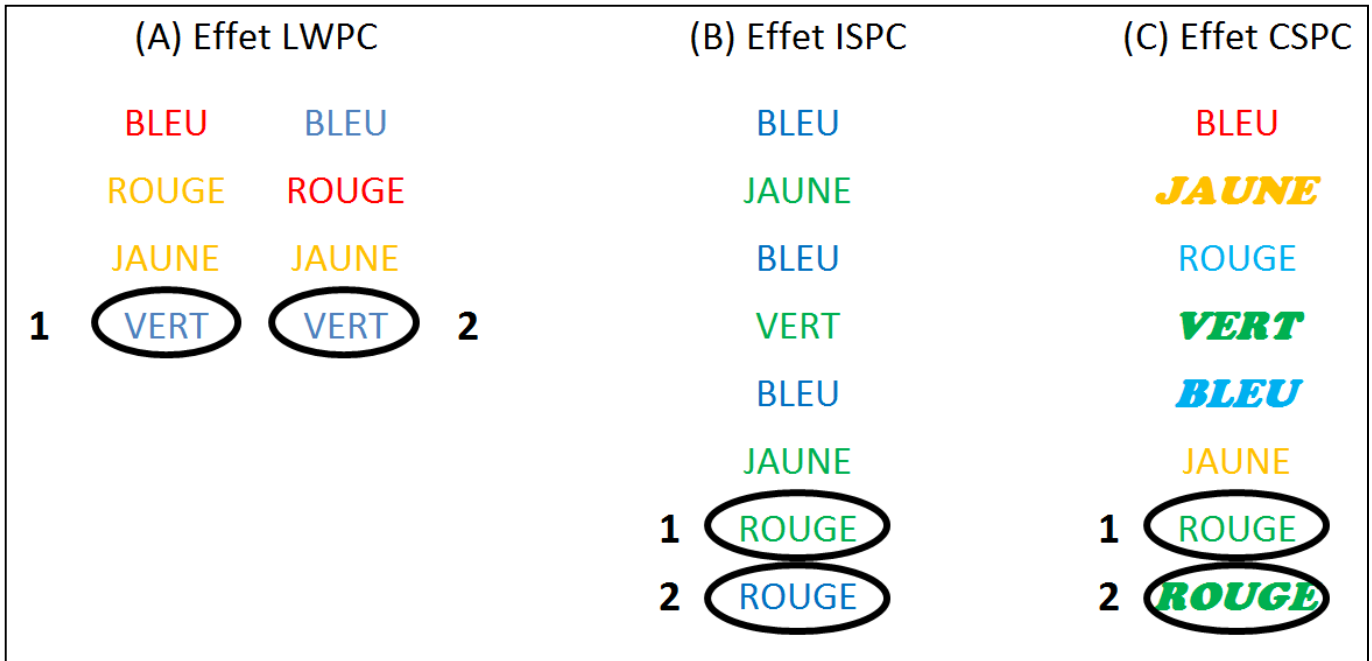
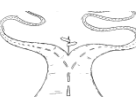


Figure 3. Illustration des trois principaux effets de congruence : (A) Effet LWPC (la liste de gauche est MI, la liste de droite est MC) ; (B) Effet ISPC (la couleur verte est MI, la couleur bleue est MC) ; (C) Effet CSPC (la police de caractères fine est MI, la police de caractères gras est MC). Pour chaque effet, la comparaison des stimuli entourés notés 1 et 2 donne lieu à un temps de réponse plus rapide pour 1 que pour 2.

Les effets LWPC et ISPC constituant logiquement des reflets des mécanismes proactif et réactif, ils présentent un intérêt pratique important puisqu'ils doivent permettre d'étudier ces deux modes de contrôle en se détachant du paradigme AX-CPT. Ce constat nous a conduits à réaliser la première étude des deux effets au sein du même échantillon, selon un plan intra-sujets, avec une variante de la tâche de Stroop (Gonthier, Braver, & Bugg, 2014, 2016). Dans cette étude, nous avons quantifié l'effet LWPC grâce à des essais 50% congruents présentés au sein de blocs MC et MI, et l'effet ISPC grâce à des essais MC et MI présentés au sein d'un bloc d'essais 50% congruent, de façon à obtenir des indicateurs du contrôle proactif et réactif respectivement. Notre objectif était quadruple : vérifier si nous arrivions à répliquer l'existence de ces effets décrits par la littérature au sein d'un même échantillon ; comparer leur magnitude ; vérifier que les résultats démontraient une signature comportementale compatible avec le fonctionnement supposé des deux mécanismes de contrôle ; et vérifier si la taille des effets LWPC et ISPC était corrélée pour un même participant, de façon à argumenter en faveur ou à l'encontre de l'indépendance des deux mécanismes.



Les résultats nous ont permis d'observer les deux effets, avec une magnitude comparable ; ils ont aussi montré un pattern de désavantages compatibles avec le fonctionnement hypothétique des deux mécanismes de contrôle, l'effet LWPC générant un ralentissement sur les essais congruents (ce qui peut correspondre à l'application sans discrimination d'un mécanisme de contrôle consistant à réduire le poids accordé à la dimension non pertinente mais prépondérante du stimulus, dont la prise en compte partielle faciliterait habituellement le traitement des stimuli congruents), et l'effet ISPC générant peu de bénéfice pour des stimuli de transfert congruents à 50% (ce qui peut correspondre à l'application sélective d'un mécanisme de contrôle consistant à réduire le poids accordé à la dimension non pertinente uniquement sur un type de stimulus particulier qui le nécessite fréquemment). L'absence de corrélation entre les deux effets a été interprétée comme un argument en faveur de l'indépendance du contrôle proactif et du contrôle réactif, un même participant pouvant présenter un fort bénéfice lorsqu'il applique un contrôle proactif et un faible bénéfice lorsqu'il applique un contrôle réactif.



La majorité des travaux concernant les effets de proportion de congruence ayant porté sur la tâche de Stroop, nous avons également cherché à étendre l'étude des effets LWPC et ISPC à d'autres paradigmes. La flanker task décrite un peu plus haut apparaissait comme un choix logique : il s'agit d'une tâche de conflit similaire dans sa composition à la tâche de Stroop, tout en étant suffisamment différente, notamment de par la nature visuo-spatiale des stimuli, pour constituer un vrai test de la généralisabilité des phénomènes en jeu. Nous avons en particulier cherché à savoir si l'effet LWPC pouvait s'observer en-dehors de la tâche de Stroop sur des items 50% congruents dans les deux listes MI et MC. En effet, les quelques travaux ayant cherché à manipuler la proportion de congruence des blocs d'essais dans la flanker task (e.g. Gratton et al., 1992 ; Lehle & Hübner, 2008 ; Wendt & Luna-Rodriguez, 2009) l'avaient fait en biaisant l'ensemble des stimuli (tous les stimuli de la liste MI étant MI, et tous les stimuli de la liste MC étant MC), ce qui ne permet pas de dissocier la contribution des mécanismes proactif et réactif puisqu'une différence entre les deux listes peut alors s'expliquer aussi bien par la contribution d'un mécanisme fonctionnant au niveau des items que par un mécanisme fonctionnant au niveau global de la liste.

Une série d'études (Bugg & Gonthier, 2020) nous a permis de montrer qu'il est effectivement possible d'observer dans la flanker task un effet LWPC sur des essais 50%

congruents dans les deux listes, et donc reflétant théoriquement un pur effet du contrôle proactif. Nous avons obtenu ces résultats sur deux versions différentes de la flanker task, une version basée sur des lettres (E E J E E) et une version basée sur des flèches (> > < > >). Conjointement avec les travaux existants qui ont montré l'existence d'un effet ISPC au sein de blocs d'essais 50% congruents (Bugg, 2015), ce travail a permis de montrer que les effets LWPC et ISPC ne sont pas spécifiques à la tâche de Stroop mais se généralisent bel et bien à d'autres tâches de conflit.

Une difficulté particulière de ce travail a consisté à trouver un design permettant d'isoler l'effet LWPC dans le contexte d'une flanker task basée sur les flèches. Les contraintes expérimentales pour l'identification de l'effet LWPC sont assez abstruses. En bref, obtenir un effet LWPC aussi dépourvu de biais que possible nécessite normalement six stimuli différents : quatre stimuli biaisés servant à manipuler la proportion de congruence, MI dans la liste MI et MC dans la liste MC, qui doivent être au nombre de quatre au minimum pour éviter l'apprentissage associatif de contingences de type "dans la liste MI je dois répondre l'inverse du stimulus présenté" ; et deux stimuli non biaisés, congruents à 50% dans les deux listes, qui permettent de tester l'effet LWPC proprement dit. Une difficulté supplémentaire est qu'aucun essai ne devrait permettre aux deux stimuli non biaisés de présenter un chevauchement perceptif avec les items biaisés MC/MI, au risque de permettre au participant de transférer un mécanisme item-spécifique aux items non biaisés. Par exemple, si les lettres ABCD servent de stimuli biaisés MC/MI et les lettres EF de stimuli non biaisés, il ne devrait pas y avoir d'essai de type AAEEA ou bien EEAE (voir Braem et al., 2019). Malheureusement, les versions habituelles de la flanker task basée sur les flèches ne comprennent que deux stimuli (< et >), tandis que certaines versions modifiées en comprennent quatre (< > v ^). Ce travail nous a permis de proposer un design modifié adapté pour tester l'existence d'un effet LWPC dans une tâche à quatre stimuli (Bugg & Gonthier, 2020), dont le principe consiste à utiliser deux items biaisés et deux items non biaisés ; à permettre un chevauchement des deux sets d'items, de façon à éviter l'apprentissage de contingences (un item non biaisé pouvant donc servir de dimension non pertinente dans le cadre d'un item biaisé) ; mais à n'analyser que les essais dans lesquels il n'y a pas de chevauchement avec le set d'items biaisés. Ce design modifié ne contrôle pas entièrement toutes les explications alternatives mais ouvre la possibilité de tester l'effet LWPC dans des tâches à quatre stimuli, ce qui permet également de le décliner à des populations pour lesquelles des tâches à six stimuli paraissent trop complexes (Bugg & Gonthier, 2020).

Dans le cadre de ces travaux sur les effets de proportion de congruence, j'ai été régulièrement insatisfait de la difficulté à s'assurer de la nature des mécanismes en jeu pendant l'implémentation proprement dite du contrôle. Nous n'observons en effet que la réduction moyenne de l'interférence entre des conditions MC et MI, sans avoir de regard sur le processus amenant à la réponse au sein de chaque essai, et même si l'identification en moyenne de signatures comportementales distinctes pour les effets LWPC et ISPC est compatible avec les hypothèses du modèle (Gonthier, Braver, & Bugg, 2016), il s'agit toujours d'un test assez indirect, et qui plus est de peu d'utilité diagnostique au niveau individuel. Ainsi, si l'effet LWPC est théoriquement causé par le même mécanisme de contrôle proactif qu'un nombre d'erreurs AY élevé sur l'AX-CPT, il reste difficile de s'en assurer, et ce d'autant plus que les deux tâches sont profondément différentes et que le déploiement des mécanismes de contrôle ne se fait pas tout-à-fait de la même façon. Dans l'AX-CPT, le contrôle proactif implique de maintenir explicitement en mémoire l'identité de l'indice pendant les quelques secondes qui précèdent l'apparition de la cible ; dans une tâche de conflit, l'effet LWPC implique d'apprendre implicitement la plus grande fréquence des essais incongruents dans un bloc d'essais donné et de maintenir un faible niveau de traitement de la dimension non pertinente pendant les plusieurs minutes que dure ce bloc d'essais.



Nous avons récemment travaillé sur un paradigme qui rend possible une analyse un peu plus fine des processus en jeu dans les effets de proportion de congruence, et qui pourrait permettre de vérifier la nature des mécanismes de contrôle qu'ils impliquent. Nous nous sommes appuyés sur une tâche de conflit de type "Stroop spatial", qui permet d'étudier plus précisément la façon dont la proportion de congruence module la dynamique de réponse (Tafuro, Vallesi, Gonthier, & Ambrosini, en révision). La tâche présentait au participant une flèche apparaissant à différents endroits de l'écran, et donnait pour consigne de cliquer sur l'emplacement correspondant à la direction indiquée par la flèche - et non sur la flèche elle-même. Cette procédure est illustrée dans la Figure 4A. L'originalité de cette tâche consiste à pouvoir analyser la trajectoire et la vitesse du curseur de la souris servant au participant à effectuer sa réponse, ce qui ouvre une fenêtre sur la dynamique temporelle de la réponse au sein même d'un essai. On peut ainsi étudier l'interférence en déterminant à quel point la réponse du participant est "attirée" par la dimension non pertinente du stimulus (la position spatiale de la flèche) avant de s'orienter vers la dimension pertinente (la direction indiquée par

la flèche). Cela permet non seulement de quantifier l'interférence, mais aussi de déterminer à quel moment de l'essai elle se déploie et à quelle vitesse elle est résolue.

La Figure 4B propose une illustration des trajectoires de réponse dans différentes conditions expérimentales, incluant des essais congruents et incongruents et manipulant la proportion de congruence (25%, 50% ou 75% congruent) au niveau d'un bloc d'essais (LW) ou d'un type d'item (IS). Ces résultats suggèrent un ensemble de conclusions. D'une part, les essais incongruents génèrent effectivement une tendance à répondre en direction de la dimension non pertinente du stimulus ; d'autre part, les effets LWPC et ISPC se manifestent sous la forme d'une correction de trajectoire plus précoce lorsque les stimuli sont généralement incongruents que lorsqu'ils sont généralement congruents. L'effet LWPC induit une correction légèrement plus précoce que l'effet ISPC (environ 30 millisecondes), ce qui semble compatible avec l'hypothèse selon laquelle l'effet LWPC représente une forme de contrôle proactif appliqué de façon anticipatoire. Ces résultats restent préliminaires, mais font apparaître le mouse-tracking comme une méthodologie prometteuse pour mieux comprendre le processus de réponse au sein d'un essai et la façon dont ce processus peut varier en fonction du mécanisme de contrôle employé.



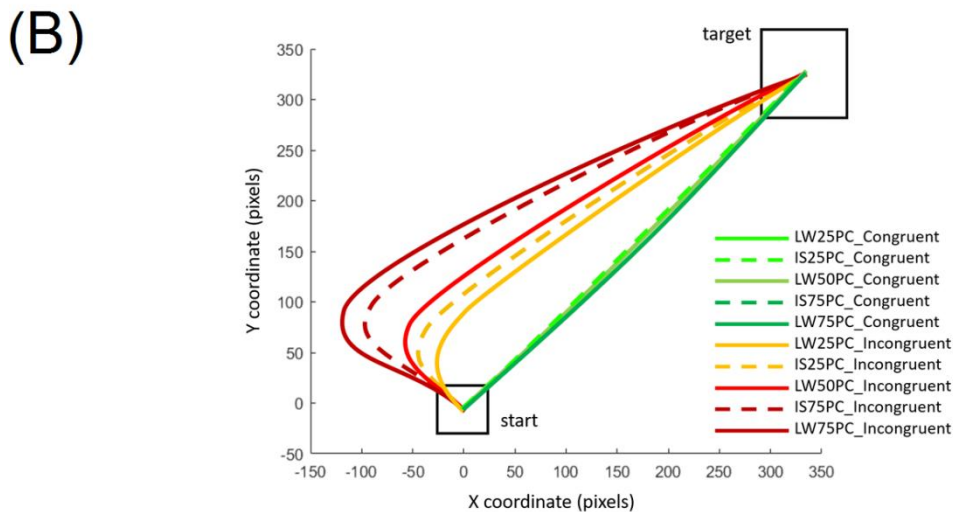
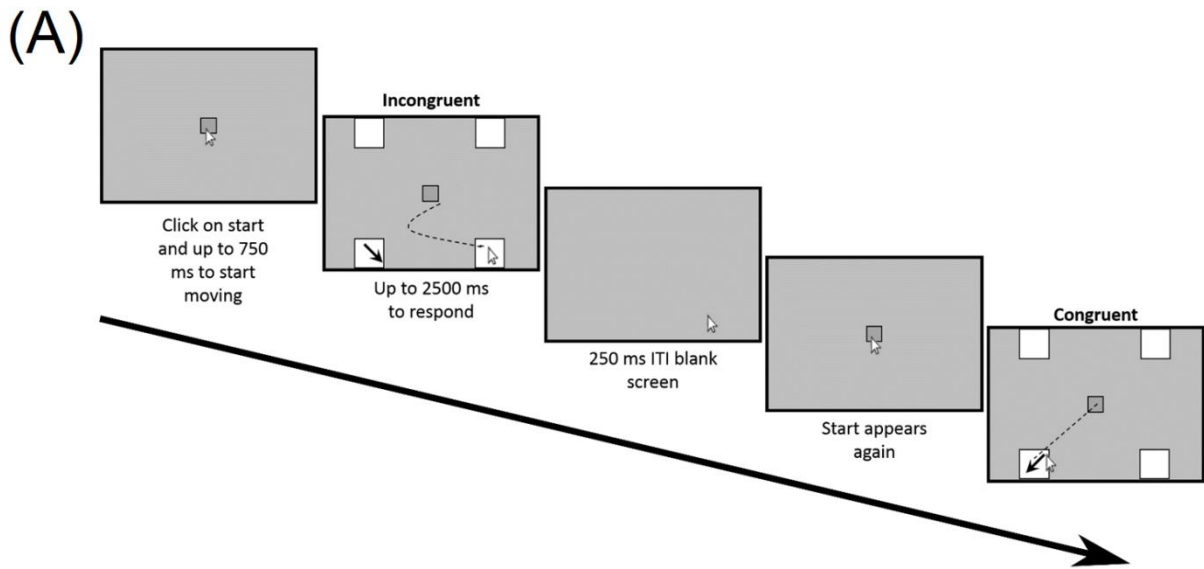


Figure 4. (A) Paradigme de Stroop spatial permettant d'évaluer la trajectoire de la réponse d'un participant ; (B) illustration des trajectoires de réponse pour des essais congruents ou incongruents, dans le cadre d'un bloc d'essais (LW) ou d'un type d'item (IS) soit 25% congruent (25PC), soit 50% congruent (50PC), soit 75% congruent (75PC). Tiré de Tafuro, A., Vallesi, A., Gonthier, C., & Ambrosini, E. (en révision). *Tracking proactive and reactive cognitive control in interference resolution*.

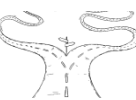
3. Étudier le développement du contrôle cognitif

3.1. Le développement du contrôle à travers l'AX-CPT

De nombreux travaux ont montré que l'aptitude au contrôle cognitif augmente avec l'âge au cours du développement (e.g. Diamond, 2013). En tant que telle, cette conclusion n'est qu'à moitié satisfaisante car elle ne nous apprend pas grand-chose sur les mécanismes en jeu : les enfants sont certes de plus en plus apte avec l'âge à réguler leur activité cognitive pour atteindre un but, mais quel aspect exactement du processus de contrôle se développe ? Est-ce leur capacité à se souvenir du but (mémoire de travail), à l'implémenter en pratique (contrôle cognitif), à s'apercevoir qu'ils ne sont pas en train de le suivre (métacognition), etc ? Les travaux cherchant à étudier séparément le développement de différentes fonctions exécutives (voir e.g. Diamond, 2013 ; Wiebe et al., 2008) n'ont pas permis d'apporter beaucoup plus d'informations à ce niveau, ni les auteurs qui invoquent sans plus de précision la "maturation du cortex préfrontal" (voir e.g. Johnson et al., 2009).



Les premiers travaux portant sur le modèle DMC ont largement étudié le vieillissement normal (e.g. Paxton et al., 2006, 2008), le passage d'un contrôle proactif à un contrôle réactif étant supposé à l'origine du déclin de la performance exécutive avec l'âge. Il était donc assez logique de chercher à utiliser ce même modèle pour comprendre l'évolution du contrôle à l'autre extrémité du spectre : dans le cadre du développement cognitif de l'enfant. La littérature s'est orientée vers une hypothèse symétrique aux résultats observés dans le cadre du vieillissement normal : l'efficacité du contrôle cognitif augmenterait avec l'âge parce que les enfants passent progressivement d'un mode de contrôle réactif à un mode de contrôle proactif. Ainsi, les jeunes enfants, avant 5 ans environ, seraient réactifs, et les enfants plus âgés seraient proactifs. On trouve dans la littérature un certain nombre d'affirmations catégoriques en ce sens (e.g. "les jeunes enfants sont exclusivement réactifs" ; Chevalier, 2015 ; Munakata et al., 2012). Nous avons donc été surpris de constater qu'aucune étude, en réalité, n'avait jamais observé directement un pattern de performance réactif chez les enfants les plus jeunes dans une tâche permettant sans ambiguïté de déterminer le type de contrôle utilisé.



Le paradigme AX-CPT permet pourtant de tester assez simplement cette hypothèse : il suffit de vérifier que les enfants les plus jeunes ont à la fois une performance plus faible en BX, et surtout une performance plus élevée en AY, qui permet de conclure avec une relative

certitude que le développement engendre bien un changement qualitatif dans le mécanisme de contrôle utilisé. Parmi les études qui ont proposé une transition développementale du contrôle réactif vers le contrôle proactif, certaines ont montré une augmentation du contrôle proactif avec l'âge, sans parvenir à observer un pattern réellement réactif (avec une meilleure performance en AY qu'en BX) même chez les enfants les plus jeunes ; ce résultat revient essentiellement à montrer une simple augmentation de performance avec l'âge, et ne donne pas d'argument fort en faveur d'un changement qualitatif dans le mode de contrôle. L'étude la plus souvent citée à l'appui d'une transition entre réactif et proactif (Chatham et al., 2009) a bien observé un pattern de dilatation pupillaire compatible avec un plus faible usage du contrôle proactif chez des enfants de trois ans, mais le pattern de taux d'erreurs et de temps de réponse ne permettait pas de conclure clairement à l'usage d'un contrôle réactif chez les plus jeunes : les taux d'erreurs AY et BX étaient uniformément proches du niveau du hasard, suggérant plutôt une difficulté à s'appropriier la tâche à cet âge. D'autres études ont tout simplement invoqué l'usage accru du contrôle proactif comme explication d'une augmentation de performance avec l'âge, sans utiliser de tâche qui permette d'évaluer quel mécanisme de contrôle était mis en jeu par l'enfant (e.g. Doebel et al., 2018).

Dans une expérience portée par Agnès Blaye (Gonthier, Zira, Colé, & Blaye, 2019), nous avons donc cherché à étudier la transition entre contrôle réactif et contrôle proactif à l'aide de l'AX-CPT chez des enfants d'âge préscolaire (4 à 7 ans), avec un triple objectif : vérifier si un pattern réellement réactif pouvait être observé chez les enfants les plus jeunes ; situer l'âge moyen de la transition entre réactif et proactif ; et étudier un déterminant possible de cette transition, la capacité en mémoire de travail. En théorie (voir page 25), l'utilisation du contrôle proactif requiert en effet la maintenance active d'informations contextuelles en mémoire de travail (le but de la tâche, ou l'identité de la cible dans le paradigme AX-CPT), et l'augmentation de la capacité en mémoire de travail chez l'enfant peut donc jouer un rôle dans l'utilisation croissante du contrôle proactif.

Nos résultats ont montré que les enfants les plus jeunes présentaient bel et bien un pattern réactif classique, avec une performance AY plus élevée que leur performance BX, la bascule entre réactif et proactif pouvant être estimée autour de 5 ans et demi en moyenne. Les résultats ont également permis de montrer que la mémoire de travail était corrélée à l'usage du contrôle proactif tel qu'indexé par le PBI, et que le lien entre âge et contrôle proactif n'était

plus significatif en contrôlant pour la mémoire de travail : autrement dit, l'augmentation du contrôle proactif avec l'âge va de pair avec l'augmentation de la capacité en mémoire de travail. Ces résultats ont permis de confirmer l'hypothèse d'une transition du contrôle réactif vers le contrôle proactif chez le jeune enfant, pouvant contribuer à l'amélioration progressive de la performance de contrôle cognitif au cours du développement. Le constat que l'usage du contrôle proactif augmente conjointement avec la capacité en mémoire de travail ne permet pas, bien sûr, de conclure avec certitude à un rôle causal de la mémoire de travail dans le développement du contrôle proactif, mais va dans le sens de cette hypothèse.

Plus généralement, ce résultat a aussi un intérêt méthodologique : comme évoqué plus haut, j'ai échoué à de nombreuses reprises dans le cadre de mon travail doctoral à observer un lien entre mémoire de travail et contrôle proactif (Gonthier, 2014), mais ces échecs peuvent être en partie mis en lien avec la faible fidélité des mesures obtenues dans l'AX-CPT lorsqu'il est utilisé chez des adultes, pour lesquels l'étendue des mesures est faible (Cooper, Gonthier, Barch, & Braver, 2017 ; voir pages 30 et 32). S'il est possible que le lien entre mémoire de travail et contrôle proactif n'existe que chez l'enfant (leur faible capacité en mémoire de travail plaçant une contrainte plus forte sur l'usage du contrôle proactif, ou l'âge étant à l'origine d'une corrélation factice), on peut donc envisager une alternative : un échantillon d'enfants d'âges variables présenterait une plus forte variabilité, aussi bien de la capacité en mémoire de travail que du contrôle cognitif, qui permettrait de mettre en évidence plus facilement la corrélation entre ces deux construits en améliorant les propriétés métriques de la tâche. À mon sens, cette possibilité illustre l'intérêt d'étudier le développement cognitif non seulement pour lui-même, mais aussi comme un contexte permettant l'évaluation de déterminants causals de la performance postulés chez l'adulte mais difficiles à mettre en évidence lorsque la performance est uniformément élevée.

3.2. Le développement du contrôle à travers les effets de proportion de congruence

Comme je l'ai évoqué, il me paraît important de chercher à généraliser l'étude des mécanismes de contrôle à d'autres tâches que le paradigme AX-CPT, pour assurer la validité du modèle DMC et des conclusions que l'on en tire. Dans le contexte de la recherche développementale en particulier, le paradigme AX-CPT n'a pas que des avantages, même lorsqu'on utilise des versions adaptées à l'enfant. On peut relever, entre autres, la grande complexité des consignes pour des enfants de cinq ans (*tu vas voir des séries de deux animaux qui font la course, on cherche une poule qui fait la course avec un chat, tu devras toujours appuyer sur "non" quand tu verras le premier animal puisque tu n'auras pas encore vu une poule faire la course avec un chat, tu devras appuyer sur "oui" si le deuxième animal est un chat mais seulement si le premier animal était une poule, et pas si le premier animal n'était pas une poule*), ou encore le caractère faiblement intéressant de la tâche qui impose un faible nombre total d'essais, et donc un faible nombre d'essais critiques AY et BX.

Les effets de proportion de congruence (LWPC et ISPC) décrits plus haut présentent un intérêt particulier dans le cadre d'une étude du développement du contrôle cognitif. D'une part, si l'on admet que ces effets évaluent respectivement le fonctionnement du contrôle proactif et du contrôle réactif, ils ouvrent la possibilité de tester le développement du contrôle proactif en s'abstrayant de l'AX-CPT. D'autre part, le fait que la nature des tâches de conflit diverge assez nettement de l'AX-CPT ouvre la possibilité d'étudier l'existence de dissociations fonctionnelles au sein même du contrôle proactif et du contrôle réactif, en s'appuyant sur la nature plutôt implicite des mécanismes mis en œuvre par les tâches classiques de conflit.

Plusieurs auteurs ont en effet argumenté que le contrôle cognitif peut être implémenté soit de façon explicite (de façon intentionnelle, tel que guidé par les attentes conscientes du participant), soit de façon implicite (sur la base d'un apprentissage implicite des régularités de la tâche, le participant étant inconscient de la régulation mise en place ; voir Hommel, 2007, 2013 ; Van Gaal et al., 2012). Ainsi, l'usage d'un contrôle proactif dans l'AX-CPT ferait plutôt appel à un contrôle explicite, le participant étant conscient de la nature de l'indice, et de la préparation et du maintien actif de la réponse qui doit s'ensuivre pendant le délai qui sépare l'indice de la cible ; à l'inverse, l'effet LWPC tel qu'il apparaît par exemple dans une tâche de Stroop ferait appel à un contrôle implicite, puisqu'il s'appuie sur l'apprentissage implicite de la



fréquence du conflit dans une condition donnée (Cohen-Shikora et al., 2018), visiblement sans que le participant en soit conscient (Blais et al., 2012). Ces deux paradigmes feraient donc appel à deux formes différentes de contrôle cognitif, déclenché de façon explicite (par les attentes du participant) ou implicite (par les régularités de la tâche).

Cette possibilité soulève une double question : y a-t-il une véritable différence fonctionnelle dans l'opération du processus de contrôle cognitif, selon qu'il est déclenché par des événements explicites ou implicites - ou bien cette distinction ne renvoie-t-elle qu'à une nuance dans les paradigmes expérimentaux qui n'a de sens que pour les chercheurs ? De façon symétrique, les enfants d'âge préscolaire ont-ils vraiment de grandes difficultés dans l'application du contrôle cognitif, lorsqu'il est recruté de façon implicite ? Les résultats obtenus avec l'AX-CPT suggèrent que les enfants les plus jeunes n'utilisent pas le contrôle proactif et ne doivent pas montrer d'effet LWPC ; étant données les difficultés fréquemment rapportées chez l'enfant à alterner explicitement entre plusieurs registres de tâche (*mental task sets* ; voir par exemple Reimers & Maylor, 2005), on peut même douter de leur capacité à démontrer un effet ISPC, qui requiert une modulation fine des paramètres de contrôle à chaque essai selon que le stimulus présenté apparaît comme MC ou MI. Observer un effet LWPC en particulier chez les enfants les plus jeunes amènerait à la fois à nuancer la trajectoire développementale du contrôle proactif, et à soutenir l'idée d'une véritable distinction fonctionnelle entre contrôle explicite et implicite, le premier étant mis en place plus tardivement que le second.

C'est l'hypothèse que nous avons testée dans une série de cinq expériences (Gonthier, Ambrosi, & Blaye, en révision), dans lesquelles nous avons vérifié l'existence des effets LWPC et ISPC chez des enfants de 5-6 ans, dans une tâche de Stroop ainsi que dans une tâche de flanker basée sur les flèches utilisant le design à quatre stimuli proposé précédemment (Bugg & Gonthier, 2020). Nos résultats ont montré des effets LWPC et ISPC importants chez les enfants de 5-6 ans, et ce dans les deux types de tâches. Plus surprenant, ces deux effets sont apparus comme plus marqués chez les enfants de 5-6 ans que dans un groupe de comparaison constitué d'enfants de 9-10 ans. Ce résultat inattendu semble pouvoir s'attribuer à la plus grande interférence observée chez les enfants les plus jeunes, qui est donc plus réduite en proportion par la mise en place du contrôle que chez les enfants plus âgés ; cette conclusion va dans le sens de l'idée évoquée plus haut selon laquelle des effets expérimentaux difficiles à observer chez l'adulte peuvent l'être plus facilement chez l'enfant. Dans le cas de l'effet LWPC, cette

trajectoire est en tout cas particulièrement incompatible avec l'augmentation du contrôle proactif observée sur l'AX-CPT et parle en faveur d'une dissociation fonctionnelle entre contrôle proactif explicite et implicite. Elle souligne ainsi l'existence de compétences précoces dans l'utilisation d'un contrôle global, soutenu, lorsque la mise en place de ce contrôle est favorisée par la situation.

Quelles sont les limites de l'aptitude des jeunes enfants à mettre en place un contrôle cognitif implicite ? Pour répondre à cette question, nous avons également souhaité nous intéresser au développement de l'effet CSPC. Théoriquement, le mode de contrôle cognitif requis pour obtenir un effet CSPC ne diffère pas de l'effet ISPC - il s'agit de réguler le poids relatif des dimensions congruente et incongruente à la volée, après l'apparition du stimulus, sur la base du degré de congruence habituel de ce type de stimulus, et il s'agit donc de contrôle réactif, dont on peut s'attendre à ce qu'il existe même chez les jeunes enfants. Mais il existe une différence majeure entre les deux effets : pour l'effet CSPC, la fréquence du conflit est indexée par une dimension qui n'a pas de lien direct avec la tâche, et à laquelle le participant n'a pas de raison particulière de prêter attention. L'effet est d'ailleurs relativement difficile à observer chez l'adulte, toujours de faible magnitude, et fonctionne mieux lorsqu'il dépend d'une dimension perceptive qui joue un rôle plus ou moins impératif dans le traitement cognitif, telle que la localisation spatiale (voir Crump et al., 2006, 2008). Observer un effet CSPC chez l'enfant soutiendrait donc l'hypothèse, non seulement qu'ils sont capables d'utiliser efficacement une forme implicite de contrôle cognitif, et que ce contrôle peut être ajusté de façon flexible à la présentation de chaque nouvel item, mais également que ce contrôle peut s'appuyer sur l'apprentissage implicite d'aspects relativement fins de la tâche qui peuvent même présenter des difficultés pour l'adulte.

Une série de deux expériences (Gonthier & Blaye, en préparation) à l'aide d'une tâche de flankers basée sur les flèches nous a permis de montrer qu'à l'image de l'effet ISPC, les enfants de 5-6 ans présentent effectivement un effet CSPC, d'une magnitude significativement plus élevée que des enfants de 9-10 ans, et au moins aussi grande que chez l'adulte. Nous avons pu mettre en évidence cet effet en manipulant aussi bien l'emplacement des stimuli, que le type de stimuli (forme perceptive de la flèche). L'observation d'un effet CSPC même sur des stimuli de transfert (des items 50% congruents présentés dans les contextes MC et MI) a permis de confirmer que cet effet n'était pas dû à un apprentissage de la réponse la plus fréquente à

donner à des stimuli particuliers apparaissant dans un contexte particulier, mais bien à une modulation dynamique du contrôle cognitif en fonction du contexte de présentation des items. Ces résultats suggèrent, à nouveau, que même les jeunes enfants s'avèrent capables d'utiliser une forme assez complexe de contrôle cognitif lorsqu'il est guidé par un apprentissage implicite des régularités de la tâche.

4. Induire l'usage d'un mécanisme de contrôle

4.1. Induire le contrôle proactif et réactif chez l'adulte

Le modèle DMC a plutôt été développé pour chercher à expliquer la variabilité spontanée du contrôle cognitif, notamment dans le cadre du vieillissement normal et de la schizophrénie. Mais quelques travaux se sont également intéressés à l'induction expérimentale de l'un ou l'autre mode de contrôle dans le cadre de l'AX-CPT : autrement dit, pousser un participant à utiliser sélectivement le contrôle proactif ou le contrôle réactif. Une telle induction présente au moins trois intérêts :

1) pouvoir évaluer le substrat neuronal qui sous-tend le fonctionnement de l'un ou l'autre mécanisme dans une tâche, préoccupation fréquente pour les chercheurs qui étudient le modèle DMC, dans la mesure où le modèle propose des hypothèses claires sur l'activité cérébrale (activité soutenue du cortex dorsolatéral préfrontal pour le contrôle proactif, activité transitoire associant également le cortex cingulaire antérieur pour le contrôle réactif) ;

2) pouvoir mesurer l'efficacité d'un mécanisme en particulier : dans son usage normal, les différences inter-individuelles dans l'AX-CPT peuvent représenter aussi bien des différences dans le type de contrôle utilisé que dans l'efficacité du contrôle utilisé ; par exemple, un indice *d'*-context plus élevé indique un taux d'erreur BX plus faible qui peut être dû autant à un usage plus important du contrôle proactif, qu'à d'autres facteurs comme une vitesse de traitement plus élevée, un meilleur rappel de l'identité de l'indice, etc ;

3) pouvoir tester des hypothèses sur l'origine causale de la variabilité. Par exemple, dans le contexte des différences entre groupes : si l'on fait l'hypothèse que les différences de contrôle cognitif observées entre personnes âgées et jeunes adultes sont avant tout dues à l'usage d'un contrôle proactif chez les jeunes et réactif chez les âgés, il doit être possible d'éliminer ces différences en induisant l'usage du contrôle réactif chez les plus jeunes. La même démarche est possible dans le contexte de l'étude des corrélations avec d'autres construits, tels que la mémoire de travail, et donne lieu dans ce cas à une démarche expérimentale-corrélacionnelle (Cronbach, 1957) : parvenir à faire disparaître la corrélation entre deux construits grâce à une manipulation expérimentale, c'est avoir l'assurance que la manipulation en question a affecté de façon causale le mécanisme à l'origine de cette corrélation. Cette démarche est particulièrement importante dans une perspective constructiviste puisqu'elle permet de



s'interroger sur le rôle des différents mécanismes en jeu dans la détermination d'un score, et donc d'une corrélation.

Les quelques travaux qui ont cherché à induire l'usage d'un mécanisme de contrôle l'ont fait dans le contexte d'autres projets (e.g. Braver et al., 2009; Edwards et al., 2010; Paxton et al., 2006, 2008), sans vraiment chercher à vérifier systématiquement l'effet de ces manipulations, ni à développer des techniques simples et qui pourraient fonctionner dans tous types de populations. C'est pourquoi nous nous sommes intéressés à développer des techniques générales d'induction du contrôle réactif ou proactif dans l'AX-CPT.

Une série de trois expériences (Gonthier, MacNamara, Chow, Conway, & Braver, 2016; Chow, Gonthier, MacNamara, Conway, & Braver, 2014; Conway, Gonthier, Macnamara, Chow, & Braver, 2017) nous a permis de proposer deux manipulations. La première, permettant d'induire un contrôle réactif, consiste à insérer de rares essais no-go matérialisés par une cible qui demande au participant de fournir aucune réponse (ni la réponse "oui" attendue pour les essais AX, ni la réponse "non" attendue pour les essais AY, BX ou BY) ; sans changer fondamentalement la nature de la tâche, ces rares essais diminuent la prévisibilité de la réponse (par exemple, le participant ne peut plus préparer avec certitude la réponse "non" lorsque l'indice n'est pas un A), et font apparaître à la fois une diminution de la performance BX et une augmentation de la performance AY. La seconde manipulation, permettant d'induire un contrôle proactif, consiste tout simplement à demander au participant de préparer sa réponse à l'avance - après avoir explicité la nature de la tâche et notamment la forte probabilité conditionnelle de voir apparaître la cible X après l'indice A, et en réalisant quelques essais d'entraînement au cours duquel le participant est invité à annoncer à voix haute la réponse probable pendant le délai séparant l'indice de la cible. Cette manipulation permet, à l'inverse, de faire diminuer la performance AY tout en augmentant les erreurs BX. Ce mode d'induction a été baptisé "entraînement stratégique au contrôle proactif", suivant les articles qui l'avaient proposé à l'origine (Paxton et al., 2006, 2008).

Les difficultés initiales que nous avons eues à mettre en évidence les effets bénéfiques de ces deux manipulations illustrent à nouveau le problème de la variabilité dans l'usage de l'AX-CPT chez de jeunes adultes sains. L'induction d'un contrôle proactif par l'intermédiaire d'un entraînement stratégique dans une version basique de la tâche n'a pas permis, à elle seule, de faire apparaître une diminution des erreurs BX (bien qu'une accélération des temps de réponse

ait été observée) ; cela s'explique facilement car les taux d'erreurs BX étaient déjà proches de zéro pour nos participants, jeunes adultes utilisant déjà beaucoup le contrôle proactif. À l'inverse, l'induction d'un contrôle réactif par l'ajout d'essais no-go a généré peu d'amélioration de la performance en AY ; mais même si les essais AY sont de loin ceux qui provoquent le plus d'erreurs chez les jeunes adultes utilisant le contrôle proactif, ce taux d'erreur reste faible (environ 10% d'erreurs en moyenne, soit une seule erreur sur une tâche d'une centaine d'essais et comptant 10% d'essais AY), laissant là encore peu de marge d'amélioration. La solution que nous avons retenue a été de combiner ces deux manipulations, en comparant une condition réactive (essais no-go) à une condition proactive (essais no-go + strategy training) ; cette approche a permis de générer un taux d'erreurs plus important dans chaque condition, et en retour, d'observer nettement le pattern attendu (voir la Figure 5 pour un exemple).

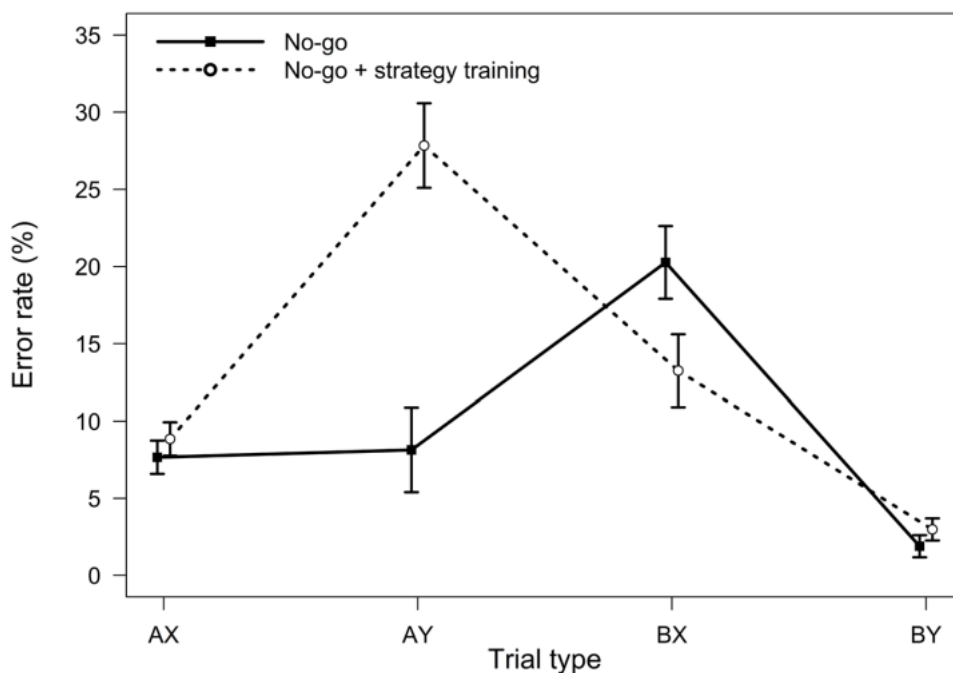


Figure 5. Taux d'erreurs dans l'AX-CPT dans deux conditions expérimentales induisant un contrôle réactif (No-go) ou un contrôle proactif (No-go + strategy training). Tiré de Gonthier, C., Macnamara, B. N., Chow, M., Conway, A. R. A., & Braver, T. S. (2016). *Frontiers in Psychology*, 7(1822).

4.2. Induire le contrôle proactif chez l'enfant ?

La possibilité d'induire l'usage d'un contrôle proactif chez l'adulte suggère une possibilité intéressante : chercher à induire également un contrôle proactif chez l'enfant. Nous avons vu que l'enfant d'âge préscolaire démontre un pattern résolument réactif dans l'AX-CPT et ne semble pas employer spontanément le contrôle proactif. D'un autre point de vue, nous avons attribué cette non-utilisation du contrôle proactif à une faible capacité en mémoire de travail, requise pour maintenir l'identité de l'indice pendant le délai qui le sépare de la cible ; mais même les enfants les plus jeunes devraient être capables de maintenir un élément en mémoire de travail pendant quelques secondes. En parallèle, nos travaux sur les effets de proportion de congruence ont montré que les enfants sont capables d'appliquer une forme de contrôle proactif lorsqu'elle est initiée par l'apprentissage implicite des régularités de la tâche. Ces deux considérations convergent pour poser la question : les jeunes enfants sont-ils incapables d'utiliser le contrôle proactif, ou bien choisissent-ils préférentiellement le contrôle réactif ?

Les déterminants de l'usage de l'un ou l'autre mode de contrôle sont en réalité mal connus. Les deux mécanismes sont parfois présentés comme deux "stratégies" de contrôle, et on retrouve cette dénomination dans l'expression "strategy training" utilisée pour désigner l'induction expérimentale d'un contrôle proactif. Ce terme semble suggérer un choix stratégique délibéré de la part des participants qui ne se justifie pas nécessairement. D'un autre point de vue, les participants sont capables d'auto-rapporter le mécanisme utilisé à l'aide d'un questionnaire (Gonthier, Macnamara, Chow, Conway, & Braver, 2016), et si l'usage croissant du contrôle réactif chez les adultes âgés est généralement attribué à une diminution de la possibilité de maintenir activement le but de la tâche au niveau du cortex dorsolatéral préfrontal, ils semblent pourtant capables de revenir à un pattern proactif lorsqu'on utilise un *strategy training* pour induire l'usage de ce mécanisme.

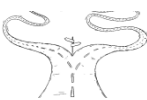
La clé se situe peut-être dans la vision classique d'une stratégie comme une procédure utilisée pour réaliser une tâche, qui ne nécessite pas d'être conscientisée, mais peut l'être dans certaines conditions (Cary & Reder, 2002 ; Lemaire & Reder, 1999). Ainsi, la sélection d'un contrôle proactif ou réactif pourrait se faire de façon non consciente, tout en pouvant être conscientisée et influencée consciemment lorsque la tâche le requiert. On pourrait donc voir les deux mécanismes de contrôle comme des processus vicariants - substituables l'un à l'autre -

faisant l'objet d'une hiérarchie d'évocabilité : les préférences personnelles du participant et ses aptitudes (sa capacité en mémoire de travail, par exemple) l'amèneraient à évoquer préférentiellement l'un des deux mécanismes, sans être pour autant incapable d'utiliser l'autre.

Dans le cas particulier des enfants d'âge préscolaire, cette idée renvoie également au modèle en *overlapping waves* de Siegler (1996, 2016) : il est concevable que les enfants aient accès conjointement aux deux mécanismes de contrôle, dont la sélection pour une tâche donnée entrerait en compétition, l'avancée en âge générant une utilisation croissante du contrôle proactif tout en laissant la possibilité d'utiliser n'importe quel mécanisme à tout âge. La littérature portant sur le développement du contrôle proactif propose quelques éléments allant dans ce sens, qui suggèrent que les jeunes enfants peuvent utiliser le contrôle proactif lorsque la situation les y encourage fortement (Chevalier et al., 2015 ; Elke & Wiebe, 2017).

Nous avons donc testé la possibilité d'induire un contrôle proactif chez l'enfant d'âge préscolaire, dans un groupe de 4-5 ans (Gonthier & Blaye, en préparation). Pour cela, nous avons utilisé la méthode du *strategy training*, qui comme nous l'avons vu consiste en réalité en un "entraînement" très simple : après avoir expliqué à l'enfant qu'un chat (X) suit souvent une poule (A), on l'engage simplement à préparer la réponse "oui" après avoir vu une poule et la réponse "non" après avoir vu un autre animal, et on lui fait expérimenter cette préparation de réponse à voix haute sur quelques dizaines d'essais. Nos résultats ont montré que les enfants de 4-5 ans placés dans cette condition expérimentale développent un pattern proactif tout-à-fait classique, similaire à celui présenté spontanément par les enfants de 6 ans et plus ; la comparaison avec une condition contrôle montre que c'est bel et bien l'induction qui a cet effet et non la simple pratique de la tâche.

Les enfants d'âge préscolaire apparaissent donc capables d'utiliser non seulement un contrôle proactif implicite, mais aussi un contrôle proactif explicite, pour peu qu'ils soient soutenus par un étayage adapté. Ce constat démontre à nouveau l'existence de compétences précoces dans l'utilisation du contrôle cognitif, qui contraste avec la vision naïve de l'enfant presque incapable de contrôler son comportement ; il amène également à repenser partiellement la variabilité développementale du contrôle cognitif, vue comme consistant non pas en l'apparition d'un nouveau mode de contrôle, mais en une évolution dans les préférences stratégiques spontanées.



5. Perspectives de recherche

5.1. Axe 1 : Trois enjeux pour le modèle DMC

L'utilisation de l'AX-CPT pour évaluer si un participant donné utilise le contrôle proactif ou le contrôle réactif laisse au moins trois problèmes majeurs en suspens.

Perspective 1.1.A : cette première perspective porte sur l'hétérogénéité méthodologique entre les implémentations de l'AX-CPT. La structure de base du paradigme est relativement constante d'une étude à l'autre, mais de nombreuses variantes existent. Dans certains cas, ces variantes ont une justification théorique. Ainsi, la première version AX-70 (70% AX, 10% AY, 10% BX, 10% BY) tend à être remplacée chez les auteurs du modèle par la version AX-40 (40% AX, 10% AY, 10% BX, 40% BY) : en diminuant la proportion d'essais AX, cette version permet de diminuer la prévisibilité des essais et d'équilibrer la fréquence des indices A et B, tout en laissant relativement inchangée la probabilité conditionnelle que la cible X suive l'indice A (0.875 dans l'AX-70 contre 0.800 dans l'AX-40). La variante DPX (MacDonald et al., 2005 ; Jones et al., 2010), de structure identique à l'AX-CPT mais qui substitue aux lettres des patterns de points tirés du braille, pourrait avoir un intérêt dans l'évaluation du contrôle cognitif dans un échantillon pour lequel l'identification des lettres n'est pas automatisée. Faire varier la longueur du délai entre indice et cible peut être un moyen de faire varier la demande en mémoire de travail (e.g. Redick & Engle, 2011).

Dans d'autres cas, les variations méthodologiques n'ont pas de justification précise et diminuent seulement la comparabilité entre études, en générant une source supplémentaire de variabilité qui peut influencer d'autres processus que le contrôle cognitif. Ainsi, dans un article combinant trois expériences réalisées dans trois universités différentes, nous avons constaté de nombreuses divergences entre nos versions de l'AX-CPT (Gonthier et al., 2016). La nature fixe ou variable des lettres B et Y est un exemple de ce type de divergences d'une étude à l'autre : dans la version originale du paradigme, "B" et "Y" représentent "toutes les lettres qui ne sont pas A et X". Mais certains auteurs utilisent uniquement les lettres B et Y (e.g. Chatham et al., 2009), ce qui diminue le statut exceptionnel de la séquence AX (il n'y a plus en effet que 4 séquences différentes dont la fréquence ne diffère que dans une proportion limitée, comparativement à la version originale dans laquelle seule la séquence AX apparaît plus d'une fois ou deux dans un bloc d'essais). D'autres auteurs utilisent "toutes les lettres sauf A et X,



moins les lettres qui présentent une similarité orthographique ou phonologique avec A ou X", ce qui pourrait retirer une autre source de variabilité non pertinente, l'aptitude à discriminer entre les lettres.

D'autres exemples de différences méthodologiques sont la présence ou non de feedbacks après chaque essai, qui pourraient moduler l'engagement dans la tâche (e.g. Mäki-Marttunen et al., 2019) ; l'usage ou non d'une limite de temps pour chaque essai, et si oui, de quelle durée (e.g. Barch et al., 2003) ; les types de feedback (classiquement "correct", "faux", ou "trop lent") et la façon dont ils se chevauchent (quel feedback pour une réponse à la fois fautive et trop lente ?) ; le fait que le découpage de la série de lettres en paires indices-cibles soit explicite ou non (Braver et al., 2005) ; dans la même logique, le fait que les listes de stimuli utilisés comme indices ou comme cibles se chevauchent ou non, et l'usage ou non d'indices visuels pour permettre au participant de savoir si c'est l'indice ou la cible qui est actuellement présenté à l'écran (par exemple en coloriant l'indice en bleu ; Henderson et al., 2012), ce qui pourrait moduler l'implication de facteurs attentionnels tels que le mind-wandering ; le fait de demander ou non une réponse à l'indice (réponse qui dans ce cas est systématiquement négative), pour s'assurer que le participant reste concentré sur la tâche et/ou pour augmenter la difficulté (Braver et al., 2001) ; les temps de présentation de l'indice et de la cible (qui varient classiquement entre 500ms et 1000ms ; e.g. Gonthier et al., 2016), qui peuvent modérer le rôle de la rapidité perceptive, mais aussi le délai entre l'indice et la cible, que la littérature fait varier entre 00500ms et 10000ms (e.g. Msetfi et al., 2009). Cette importante hétérogénéité peut certainement affecter l'usage de l'AX-CPT dans les études portant sur la variabilité, et justifie la réalisation d'un travail de revue méthodologique et de proposition d'une version standardisée.

Perspective 1.1.B : je crois impératif de faire une distinction entre efficacité du contrôle, et mode de contrôle employé. Ces deux notions sont logiquement distinctes. D'une part, je peux utiliser le contrôle proactif ou réactif ; d'autre part, je peux utiliser le contrôle proactif efficacement ou non (par exemple si j'ai mal estimé la proportion d'essais AX et que je prépare la mauvaise réponse, ou si j'oublie l'identité de la cible), et je peux utiliser le contrôle réactif efficacement ou non (par exemple si la cible X évoque fortement chez moi une tendance à donner la réponse "oui", générant une difficulté à aller récupérer l'identité de l'indice en mémoire). La différence qualitative (quel mécanisme de contrôle est utilisé ?) est pourtant

systematiquement confondue avec la différence quantitative (avec quelle efficacité je l'utilise ?) dans la littérature.

Cette confusion est en partie excusable : on a plutôt envie de tirer des conclusions qualitatives qui s'appuient sur deux mécanismes bien identifiés et permettent une explication simple à la variabilité ; la distinction entre sélection et efficacité du mécanisme n'apparaît pas dans la définition du modèle ; et l'AX-CPT n'est pas du tout construit pour l'étudier. La nature du mécanisme de contrôle que j'emploie est en effet reflétée par l'équilibre entre essais AY et BX : une plus faible performance en AY est interprétée comme une plus forte utilisation du contrôle proactif, mais pourrait en fait correspondre à une plus faible efficacité dans l'utilisation du contrôle proactif - ou même réactif, le participant ayant du mal à s'abstenir de donner la réponse "oui" qu'il avait préparée lorsque la cible s'avère être un Y et non un X.

Cette question en suspens complique fortement l'interprétation des patterns comportementaux, et ce en particulier dans l'étude des différences inter-individuelles. À mon sens, la meilleure illustration en sont les études sur le lien entre mémoire de travail et contrôle proactif : ces études trouvent systématiquement une corrélation entre mémoire de travail et performance AX, BX et BY, mais pas AY (Redick, 2014 ; Redick & Engle, 2011 ; Richmond et al., 2015). Les auteurs l'interprètent comme un plus fort usage du contrôle proactif par les participants à forte capacité en mémoire de travail - puisqu'ils sont meilleurs en BX, c'est qu'ils sont plus proactifs, et s'ils ne sont pas meilleurs en AY malgré leurs performances cognitives généralement supérieures, c'est parce qu'ils ont plus de difficultés sur ce type d'essais du fait de leur usage d'un contrôle proactif. Mais une explication alternative est que les participants à forte capacité en mémoire de travail sont tout simplement *meilleurs* dans l'utilisation du contrôle proactif, ce qui les rend ainsi généralement plus rapides, sans utiliser un mécanisme qualitativement différent des participants à faible capacité en mémoire de travail. Il existe de nombreuses raisons de penser que les participants à forte capacité en mémoire de travail pourraient utiliser ce mécanisme plus efficacement, par exemple une vitesse de traitement plus élevée (Ackerman et al., 2002 ; Fry & Hale, 2000), ou une plus grande facilité à récupérer l'indice en mémoire (Unsworth & Engle, 2007a).



Le modèle DMC est de plus en plus utilisé dans des études relevant de la psychologie différentielle, ce qui justifie à mon avis un examen direct de cette question. Elle sera pourtant difficile à traiter. Une démonstration possible passerait par l'induction de l'un ou l'autre

mécanisme, dans le cadre d'une démarche expérimentale-corrélationnelle : si une variabilité inter-individuelle subsiste, en particulier en lien avec la mémoire de travail, alors que tous les participants sont amenés à utiliser qualitativement le même mode de contrôle, c'est que la tâche est également sensible à l'efficacité générale. Nous avons débuté un projet de recherche dans cette direction, qui suggère que c'est bien ce phénomène qui se produit (Rosales, Snijder, Conway, & Gonthier, en préparation). Un autre mode de démonstration possible passerait par l'auto-report des mécanismes utilisés, les participants indiquant directement à quel point ils ont utilisé l'un ou l'autre mécanisme (ce qui permettra de vérifier si une variabilité inter-individuelle subsiste, à usage égal du contrôle proactif ou réactif). L'une de nos expériences a montré qu'un tel auto-report était possible (Gonthier et al., 2016). On peut l'envisager par exemple sous la forme de questions interrogeant la stratégie utilisée à la fin d'un essai, ou encore sous la forme d'un auto-report des pensées pendant le délai entre indice et cible (êtes-vous en train de préparer une réponse ou non ?).



Perspective 1.1.C : en lien avec la question précédente se pose le problème de l'indépendance entre les mécanismes de contrôle. Les premiers travaux sur le modèle DMC faisaient déjà l'hypothèse que contrôle proactif et contrôle réactif puissent fonctionner de façon indépendante - un même participant pouvant utiliser les deux formes de contrôle dans le cadre d'un même essai (anticiper l'apparition de la cible en préparant une réponse, puis récupérer l'identité de l'indice en mémoire lorsque la cible apparaît). La littérature commence à s'orienter dans cette direction, avec une étude en particulier qui indique avoir réussi à montrer une indépendance entre les deux formes de contrôle grâce à une induction expérimentale (Mäki-Marttunen et al., 2019). Cette étude a fait varier à la fois la présence d'une récompense (supposée induire un contrôle proactif), et le nombre de lettres demandant une réponse de type A (1, 2 ou 3 lettres ; l'augmentation du nombre d'indices de type A étant supposée induire un contrôle réactif puisqu'elle augmente la difficulté à préparer sa réponse à l'avance). Les auteurs observant que la condition (récompense + nombre d'indices élevé), qui combine les deux inductions, génère un taux d'erreur élevé à la fois en AY et en BX, ils en concluent que les deux mécanismes sont indépendants et ont des effets additifs.



Évidemment, cette conclusion pose le problème soulevé dans le point précédent : elle mélange l'usage du contrôle et l'efficacité du contrôle. Il est normal que des participants qui sont récompensés pour avoir préparé leur réponse sur la base de trois indices différents dont

l'identité doit être maintenue en mémoire (parmi 19 lettres possibles) commettent beaucoup d'erreurs à la fois en AY (reflétant la préparation à l'avance) et en BX (reflétant la difficulté de réaliser cette préparation correctement). Plus généralement, cette question de l'indépendance ne peut pas être détachée de la différence entre sélection et efficacité : parle-t-on d'indépendance en tant que la possibilité d'utiliser les deux mécanismes simultanément, ou en tant qu'indépendance dans l'efficacité de l'usage des deux mécanismes ?

Répondre à la question de l'indépendance des deux mécanismes sera également un travail délicat. La solution de l'auto-report du mécanisme utilisé pourra être employée pour vérifier s'il existe une corrélation négative entre usage du contrôle proactif et du contrôle réactif au niveau d'un même essai (dans la logique des travaux étudiant les changements de stratégie en cours d'essai ; Ardiale & Lemaire, 2013). Une autre option consiste à explorer les effets de proportion de congruence, qui offrent un angle d'attaque différent de l'AX-CPT. Nous avons partiellement contribué à cette question en observant une corrélation nulle entre les effets LWPC et ISPC, ce que notre article présentait comme un argument en faveur d'une indépendance de l'efficacité des deux mécanismes de contrôle (Gonthier, Braver, & Bugg, 2016). Mais la validité de cette conclusion est en réalité très limitée par la fidélité catastrophique de nos mesures des deux effets : environ .50 pour les indices reflétant l'interférence, entre .00 et .30 pour les indices reflétant l'effet LWPC et l'effet ISPC. Un argument plus fort, à mon avis, est dissimulé dans nos résultats : on observe un effet ISPC au sein des différentes conditions LWPC (autrement dit, un item MC génère plus d'interférence qu'un item MI, qu'il soit présenté au sein d'un bloc MC ou MI), ce qui indique que les proportions de congruence sont apprises et utilisées pour moduler l'interférence à la fois au niveau local d'un item et au niveau global d'un bloc. Mais rien ne dit que ce résultat s'étendrait à l'AX-CPT.

5.2. Axe 2 : Dynamiques du contrôle dans les tâches de conflit et les effets de proportion de congruence

Mes travaux sur la façon dont l'interférence émerge dans les tâches de contrôle suggèrent trois extensions possibles.

Perspective 1.2.A : mes travaux portant sur l'identification des signatures comportementales du contrôle proactif et du contrôle réactif sous la forme des effets LWPC et ISPC, respectivement, posent la question de la répliquabilité et de la généralisabilité des résultats. Les coûts observés en association avec les deux mécanismes (voir page 35) se jouent à quelques dizaines de millisecondes, une taille d'effet suffisamment faible pour être instable. En parallèle, le test de Stroop n'est que l'une des tâches de conflit possibles, et la version que nous avons utilisée pour ce paradigme en est une implémentation assez particulière (Stroop mot-image, qui demande d'identifier un dessin d'animal sans tenir compte du nom d'animal qui lui est superposé ; cela permet théoriquement de diminuer le rôle joué par l'apprentissage de contingences ; Bugg et al., 2011). Si les effets LWPC et ISPC doivent être interprétés sans ambiguïté comme des manifestations du contrôle proactif et du contrôle réactif, il apparaît nécessaire de répliquer nos résultats dans un autre type de tâche. La flanker task apparaît comme un candidat idéal, puisque c'est le seul paradigme dans lequel l'existence de ces deux effets a pour l'instant été démontrée.

Perspective 1.2.B : de récentes recherches ont permis de spécifier un peu plus précisément l'origine de l'interférence produite dans une tâche de conflit de type Stroop (Augustinova et al., 2018 ; Ferrand et al., 2019), en la décomposant en trois sources. La première est le conflit dû à la tâche, généré par la coexistence de deux tâches possibles, dénommer la couleur ou lire le mot (on peut le mettre en évidence en comparant les temps de réponse pour dénommer la couleur d'un stimulus illisible, **XXXX**, ou lisible, **manuscrit**). La seconde est le conflit sémantique, généré par l'incompatibilité sémantique entre la dimension congruente et la dimension incongruente (on peut le mettre en évidence en comparant les temps de réponse pour dénommer la couleur d'un mot sans incompatibilité sémantique avec la couleur de l'encre, **manuscrit**, ou avec incompatibilité, **tomate**). La troisième est le conflit dû à la réponse, généré par le fait que la dimension non pertinente est associée à une réponse différente de la dimension pertinente (on peut la mettre en évidence en comparant les temps de réponse pour



dénommer la couleur d'un mot qui n'est pas directement associé à une réponse de couleur, **tomate**, ou d'un mot qui l'est, **rouge**).

L'existence de plusieurs sources d'interférence pose la question du locus des effets LWPC et ISPC observés dans la littérature. Ces deux effets reflètent une réduction de l'interférence en condition MI ; mais quelle source d'interférence précisément est minimisée par l'apprentissage implicite du fait que certains stimuli sont généralement incongruents ? Toutes les hypothèses sont possibles : une moindre attention portée au caractère lisible du mot qui diminue le conflit lié à la tâche, une minimisation du traitement sémantique du mot, ou de façon peut-être moins plausible, une plus faible tendance à produire la réponse associée au mot. Le paradigme de comparaison entre quatre comparaisons expérimentales (Augustinova et al., 2018 ; Ferrand et al., 2019) se prête à l'examen des effets de proportion de congruence, bien qu'il pose quelques questions pratiques (dans la mesure où les stimuli incongruents, au sens classique, ne représentent que l'une des quatre conditions expérimentales entrelacées dans la tâche). Utiliser ce paradigme pour étudier l'origine des effets de proportion de congruence permettrait de progresser dans la compréhension de ces effets identifiés depuis longtemps (Logan & Zbrodoff, 1979), mais dont la nature est toujours débattue, et pourquoi pas d'apporter un regard original sur les différences de fonctionnement entre contrôle proactif et réactif.

Perspective 1.2.C : notre utilisation d'une variante de la flanker task basée sur les flèches avec quatre modalités différentes ($\langle \wedge \vee \rangle$; Bugg & Gonthier, 2020) a soulevé la question d'une possible différence d'interférence en fonction de l'écart entre la dimension cible et la dimension interférente. Un essai ($\langle \langle \rangle \langle \rangle$), dans lequel les distracteurs sont orientés à l'opposé de la cible, génère-t-il autant d'interférence qu'un essai ($\langle \wedge \wedge \rangle \langle \wedge \wedge \rangle$) dans lequel les distracteurs sont orientés à 90° - ou plus, ou moins ? On peut faire l'hypothèse que plus les distracteurs sont incompatibles avec la cible, plus l'interférence générée est importante. Cette idée amènerait à repenser la notion "d'incongruence", habituellement considérée en tout-ou-rien, comme un "degré d'incongruence". Démontrer l'existence d'un gradient d'interférence en fonction du degré d'incongruence permettrait, en retour, d'insister sur le rôle du conflit sémantique dans la production de l'interférence - car un tel gradient ne pourrait s'expliquer ni par le conflit de tâche, ni par le conflit de réponse.

Des questions similaires ont déjà été étudiées dans la tâche de Stroop (e.g. Klopfer, 1996 ; voir aussi De Marchis et al., 2013), mais les couleurs ne se prêtent pas parfaitement à l'étude de cette question car on ne peut pas faire varier la dimension non pertinente (le mot) sur un gradient ; on peut manipuler la dimension pertinente (la couleur de l'encre) de façon continue mais pas sans générer des ambiguïtés perceptives dans la réponse à apporter. En revanche, les tâches de conflit visuo-spatiales telles que la flanker task se prêtent particulièrement bien à ce type de questionnement car les dimensions perceptives congruente et incongruente peuvent toutes deux être manipulées de façon continue. En plus de la flanker task, on peut évaluer une hypothèse similaire sur la tâche de Simon, le Stroop spatial (indiquer la direction d'une flèche, sans tenir compte du fait qu'elle apparaisse à gauche ou à droite de l'écran : on peut faire varier de façon continue l'emplacement de la flèche sur un axe gauche/droite), ou encore les tâches de comparaison de magnitude (par exemple indiquer lequel de deux chiffres présentés visuellement est le plus grand numériquement, sans tenir compte de sa taille physique à l'écran : on peut faire varier de façon continue la taille physique).

5.3. Axe 3 : Variabilité du contrôle cognitif

Perspective 1.3.A : le paradigme de mouse-tracking sur lequel nous avons travaillé dans le cadre de tâches de conflit (Tafuro et al., en préparation) ouvre des possibilités originales quant à l'étude des mécanismes du contrôle cognitif. Une application possible est l'étude de la façon dont la résistance au conflit cognitif se développe chez l'enfant. Les enfants plus âgés sont plus rapides dans l'absolu ; ils semblent aussi moins sensibles au conflit, dans le sens où leurs scores d'interférence sont plus faibles, mais il n'est pas toujours aisé de déterminer si cette interférence plus faible est due à une simple réduction en proportion de leurs temps de réponse plus rapides, ou à un contrôle cognitif véritablement plus efficace avec l'âge. L'analyse des trajectoires de réponse permet d'envisager des éléments de réponse à cette question. On peut imaginer que les enfants les plus jeunes suivent plus longtemps la trajectoire de réponse en direction de la réponse incorrecte, et qu'ils effectuent une correction plus tardive ; que les enfants plus âgés aient moins de biais initial en direction de la réponse incorrecte, dès le début du processus de réponse ; ou encore que les enfants plus âgés mettent plus longtemps avant d'initier leur réponse (des travaux préliminaires en ce sens ont été menés par Erb et al., 2016, 2017, 2018).



Des hypothèses similaires sont possibles pour étudier d'autres formes de variabilité, dans le cadre du vieillissement par exemple ; mais la généralisation la plus intéressante sera peut-être celle qui étudiera d'autres tâches de contrôle cognitif. Un exemple possible est le Trail Making Test, évoqué plus haut comme un exemple de tâche exécutive pour laquelle les mécanismes à l'origine de la performance restent flous : une difficulté à alterner entre chiffres et lettres dans le cadre du vieillissement est-elle due à une difficulté d'initiation de la réponse, une difficulté dans la recherche visuelle de la prochaine cible, une tendance à négliger l'alternance lettres-chiffres et à s'orienter vers la mauvaise réponse ? Une analyse fine des trajectoires de réponse permettrait d'évaluer dans quelle mesure une difficulté à cette tâche représente un véritable problème de "flexibilité mentale", c'est-à-dire d'alternance entre lettres et chiffres. Si la trajectoire s'oriente initialement vers le mauvais stimulus et se voit corrigée plus tardivement chez la personne âgée, la flexibilité mentale est en jeu ; si la trajectoire n'est pas plus déviée que chez un participant plus jeune et que l'exécution est simplement plus lente, alors la façon dont on interprète la performance à cette tâche est erronée.

Perspective 1.3.B : dans le contexte particulier de l'AX-CPT, la possibilité d'induire un contrôle proactif chez l'enfant d'âge préscolaire, que nous avons montrée avec Agnès Blaye, soulève la question des déterminants de l'évolution développementale du contrôle proactif. Si les jeunes enfants étaient incapables d'utiliser le contrôle proactif, l'explication pourrait être une limitation cognitive, peut-être au niveau de la mémoire de travail ; mais il semble plutôt qu'ils n'utilisent pas spontanément ce mécanisme de contrôle, tout en étant capables. Ce schéma semble relativement universel puisque nous n'avons observé aucun enfant d'âge préscolaire avec un pattern nettement proactif (Gonthier et al., 2019). Qu'est-ce qui, dans le développement cognitif, peut provoquer une contrainte suffisante sur le mécanisme de contrôle employé pour créer un pattern aussi homogène, tout en étant suffisamment malléable pour que donner aux enfants la consigne de préparer leur réponse à l'avance suffise à faire émerger un pattern proactif ?



Plusieurs candidats sont possibles. Il est possible que les jeunes enfants ne disposent pas initialement du contrôle proactif dans leur répertoire comportemental, et que l'induction du contrôle proactif leur révèle cette stratégie possible. Dans ce cas, expliquer aux enfants le principe du contrôle proactif, sans leur demander de l'utiliser, pourrait suffire à le faire émerger même chez les plus jeunes. Il est également possible que les enfants ne perçoivent pas initialement l'intérêt de ce mécanisme de contrôle (l'induction du contrôle proactif inclut un aspect métacognitif qui explicite directement la pertinence de préparer sa réponse à l'avance) ; ou bien que la préparation soit perçue comme trop coûteuse en efforts, le développement cognitif allant dans ce cas de pair avec une moindre aversion à l'effort ou un besoin de contrôle plus important. Une autre explication plausible est que les enfants jugent leur performance comme étant suffisante avec un contrôle réactif ; dans ce cas, le développement irait de pair avec une amélioration des aptitudes métacognitives dans l'auto-estimation de la performance, qui créerait le besoin d'un changement de stratégie. Tous ces déterminants possibles de l'augmentation développementale du contrôle cognitif devraient pouvoir être étudiés par l'intermédiaire de manipulations expérimentales (qui pourraient de la même façon aider à révéler les mécanismes en jeu chez le jeune adulte, puisque l'usage quasi-universel du contrôle proactif dans cette population se prête mal à l'étude de corrélations).

6. Publications et communications associées

Articles publiés et en révision

Bugg, J., & **Gonthier, C.** (2020). List-level control in the flanker task. *Quarterly Journal of Experimental Psychology*, 73(9), 1444-1459. doi:10.1177/1747021820912477

Cooper, S. R., **Gonthier, C.**, Barch, D. M., & Braver, T. S. (2017). The role of psychometrics in individual differences research in cognition: A case study of the AX-CPT. *Frontiers in Psychology*, 8(1482). doi:10.3389/fpsyg.2017.01482

Gonthier, C., Ambrosi, S., & Blaye, A. (en révision). Learning-based before intentional cognitive control: Developmental evidence for a dissociation between implicit and explicit control.

Gonthier, C., Braver, T. S., & Bugg, J. M. (2016). Dissociating proactive and reactive control in the Stroop task. *Memory and Cognition*, 44(5), 778-788. doi:10.3758/s13421-016-0591-1

Gonthier, C., Macnamara, B., Chow, M., Conway, A. R. A., & Braver, T. S. (2016). Inducing proactive and reactive control shifts in the AX-CPT. *Frontiers in Psychology*, 7(1822). doi:10.3389/fpsyg.2016.01822

Gonthier, C., Zira, M., Colé, P., & Blaye, A. (2019). Evidencing the developmental shift from reactive to proactive control in early childhood, and its relationship to working memory. *Journal of Experimental Child Psychology*, 177, 1-16. doi:10.1016/j.jecp.2018.07.001

Tafuro, A., Vallesi, A., **Gonthier, C.**, & Ambrosini, E. (en révision). Tracking proactive and reactive cognitive control in interference resolution.

Articles en préparation

Gonthier, C., & Blaye, A. (en préparation). Rethinking interference: A continuous gradient of interference by increasingly incongruent stimuli.

Gonthier, C., & Blaye, A. (en préparation). Preschoolers can be trained to use proactive control.

Gonthier, C., & Blaye, A. (en préparation). Development of context-specific proportion congruency effects.

Rosales, K. P., Snijder, J.-P., Conway, A. R. A., & **Gonthier, C.** (en préparation). Working memory capacity and dual mechanisms of cognitive control: A combined experimental/correlational approach.

Communications

Chow, M., **Gonthier, C.**, MacNamara, B., Conway, A. R. A., & Braver, T. S. (2014, Novembre – communication affichée). *Inducing proactive and reactive control shifts in the AX-CPT*. Psychonomics 2014, Los Angeles, US.

Conway, A. R. A., **Gonthier, C.**, Macnamara, B. N., Chow, M., & Braver, T. S. (2017, Janvier – communication orale). *Dual mechanisms of cognitive control: Experimentally induced shifts in proactive and reactive control in the AX-CPT*. SARMAC XII, Sydney, Australie.

Gonthier, C. (2017, Septembre – communication orale). *Un contrôle exécutif sans fonctions exécutives : La variabilité selon le modèle à deux mécanismes de contrôle (DMC)*. Séminaire invité, LPC, Aix-Marseille Université, Marseille, France.

Gonthier, C., Braver, T. S., & Bugg, J. M. (2014, Novembre – communication affichée). *Doubly dissociating proactive and reactive control in the Stroop task*. Psychonomics 2014, Los Angeles, US.

Thèse de doctorat

Gonthier, C. (2014). *Cognitive control in working memory: An individual differences approach based on the Dual Mechanisms of Control framework* [Thèse de doctorat non publiée]. Département de Psychologie, Université de Grenoble.

CHAPITRE 2. MÉMOIRE DE TRAVAIL

Résumé

L'histoire de la recherche sur la mémoire de travail suit un schéma assez similaire à celle du construit cognitif : des premières conceptualisations relativement abstraites, suivies de travaux mettant l'accent sur une conception structurelle de la mémoire de travail, vue comme résultant de l'action de plusieurs sous-systèmes ; et enfin l'émergence de modèles plus fonctionnels, qui s'intéressent aux processus de maintien de l'information en mémoire et à la façon dont la performance finale découle de ces mécanismes.

Mes travaux sur la mémoire de travail se sont répartis en deux grandes préoccupations : 1. développer des outils permettant d'évaluer la capacité en mémoire de travail en tenant compte de la multiplicité des déterminants en jeu dans la performance ; et 2. spécifier la nature des différents mécanismes intervenant dans le processus de mémorisation et pouvant, éventuellement, expliquer les corrélations observées entre la mémoire de travail et les autres tâches cognitives complexes.

1. Revue de littérature

La mémoire de travail est le système cognitif qui permet de maintenir l'information en mémoire pendant une courte durée, tout en manipulant conjointement de l'information (Baddeley & Hitch, 1974). Elle est fréquemment confondue avec la mémoire à court terme, dont la définition se restreint au maintien en mémoire de l'information pendant une courte durée, sans manipulation concurrente de l'information. Cette distinction est pourtant au cœur de la notion de mémoire de travail et du rôle que lui accorde la littérature. Le concept de mémoire de travail s'est en effet développé à partir des travaux sur la mémoire à court terme,

pour rendre compte de la nécessité de stocker temporairement l'information tout en la manipulant dans le cadre de la réalisation d'opérations cognitives complexes (Baddeley & Hitch, 1974).

En substance, on peut supposer que pour interagir avec des stimuli complexes (dans le cadre d'une tâche de raisonnement par exemple), il est nécessaire de garder en mémoire les différents aspects de ces stimuli pour pouvoir les manipuler de façon à produire une réponse. Ainsi, résoudre un syllogisme demanderait de conserver les trois propositions en mémoire simultanément, tout en les manipulant pour les combiner (Copeland & Radvansky, 2004). Cette nécessité est rendue particulièrement apparente dans le développement de systèmes informatiques experts, qui requièrent explicitement un système de maintien temporaire de l'information à traiter ; ce qui explique que le terme de *mémoire de travail* soit apparemment issu de l'informatique (Newell & Simon, 1956 ; voir Adams et al., 2018).

Cette vision de la mémoire de travail la place de fait au cœur du processus de traitement de l'information dans le cadre des activités cognitives de haut niveau. En psychologie, les premiers modèles se sont appuyés sur cette idée d'une mémoire de travail constituant l'espace de travail central de la cognition. Par exemple, Miller et collaborateurs (1960) abordent la mémoire de travail comme le système dédié au maintien actif en mémoire de la hiérarchie de buts et sous-butts à suivre pour réaliser une tâche complexe. Baddeley & Hitch (1974), dans la première version de leur modèle, cherchent explicitement à mettre en évidence le rôle du maintien temporaire de l'information dans les tâches cognitives complexes ; c'est leur insatisfaction face au peu d'éléments en faveur d'un rôle de la mémoire à court terme qui les pousse à proposer l'existence d'un système distinct de mémoire de travail.

Pour mettre en évidence ce rôle central de la mémoire de travail, une grande partie de la littérature s'est appuyée sur l'existence de différences inter-individuelles marquées et stables dans la capacité en mémoire de travail, pour rechercher des corrélations avec d'autres mesures. Et on peut dire qu'elle en a trouvé. Au premier plan se trouve la corrélation entre mémoire de travail et intelligence, qu'une méta-analyse estime à $r = .36$ (Ackerman et al., 2005) ; la littérature qualifie d'ailleurs la mémoire de travail de "meilleur prédicteur de l'intelligence" (Oberauer et al., 2005), et il a régulièrement été proposé que la mémoire de travail puisse être à l'origine du facteur général de l'intelligence (Kyllonen & Christal, 1990 ; Blair, 2006). À un niveau plus restreint, de nombreuses études ont montré que la capacité en

mémoire de travail est un bon prédicteur des aptitudes verbales (e.g. Daneman & Merikle, 1996 ; Gathercole & Baddeley, 1993 ; Linck, Osthus, Koeth, & Bunting, 2013) et numériques (Wiley & Jarosz, 2012), ou encore de la réussite académique (e.g. Gathercole et al., 2004).

Par ailleurs, une collection éclectique de recherches a montré des corrélations entre la capacité en mémoire de travail, et... à peu près tout - régulation émotionnelle (Schmeichel et al., 2010), précision du tir avec un pistolet (G. Wood et al., 2016), apprentissage d'un langage informatique (Lehrer et al., 1988), performance au basketball (Buszard et al., 2017), écriture d'essais littéraires (Benton et al., 1984), suppression des pensées intrusives (Geraerts et al., 2007), modération des effets de la dépression (Noreen et al., 2019), ou encore qualité de l'improvisation au violoncelle (De Dreu et al., 2012) ; sans compter les travaux en sens inverse qui proposent que la capacité en mémoire de travail est augmentée par le fait de pratiquer l'aérobic (Pontifex et al., 2009), courir pieds nus (Alloway et al., 2016), ou mâcher du chewing-gum (Wilkinson et al., 2002). Ce pattern très général de corrélations peut être vu, au choix, comme un argument en faveur de la centralité de la mémoire de travail pour l'activité cognitive (Engle & Kane, 2004), ou comme un sérieux problème de validité divergente dans la constitution d'un réseau nomologique cohérent (Cronbach & Meehl, 1955).



Les premiers modèles de la mémoire de travail se sont plutôt intéressés à son aspect structurel : ils conçoivent plutôt le maintien en mémoire de travail comme le produit de l'activité de quelques composantes statiques, et ont cherché à identifier ces composantes. L'exemple le plus connu est bien sûr l'incontournable modèle de Baddeley (1974, 1986, 2000), qui voit la mémoire de travail comme émergeant de l'activité de différents systèmes esclaves qui maintiennent l'information (boucle phonologique, calepin visuo-spatial, buffer épisodique) en fonction de sa nature (verbale, visuo-spatiale, ou multidimensionnelle et éventuellement en lien avec la mémoire à long terme), et d'un administrateur central, qui régule l'activité de ces systèmes esclaves et manipule l'information. D'autres exemples sont le modèle de la croix de Malte de Broadbent (1984), qui représente la mémoire de travail comme un espace de stockage en interaction avec un système de traitement, lui-même en lien avec l'activité perceptive, motrice, et avec la mémoire à long terme ; ou encore le modèle d'Atkinson & Shiffrin (1968), dont le système de "stockage à court terme" se place à l'interface entre registres sensoriels et mémoire à long terme, et assure également des fonctions de contrôle et de manipulation de l'information qui le rapprochent de la mémoire de travail.

Ce type de modèles structurels met plutôt l'accent sur le fractionnement de la mémoire de travail (et de la cognition en général) en plusieurs sous-composantes distinctes, en utilisant notamment la méthodologie de la double tâche (e.g. Baddeley, 1986 : si réaliser deux tâches simultanément ne produit pas d'interférence, alors ces deux tâches font appel à deux sous-composantes différentes), et l'étude de cas pathologiques qui permet de mettre en évidence des dissociations (e.g. Baddeley et al., 1997). Les auteurs de ces modèles consacrent relativement peu d'intérêt au fonctionnement pratique des sous-composantes, à leurs interactions, ou encore à la variabilité de ce fonctionnement. Ces thématiques ne sont pas absentes des travaux correspondants (par exemple, Atkinson et Shiffrin 1968 consacrent une discussion au rôle des stratégies de chunking, de regroupement ou d'organisation du matériel à mémoriser ; et le modèle de Baddeley met l'accent sur le rôle de l'auto-répétition subvocale pour limiter le déclin de la trace mnésique, et s'interroge parfois sur des aspects plus fonctionnels - voir notamment Baddeley, 2003), mais ne constituent pas leur principale préoccupation théorique.

Autrement dit, ces modèles s'intéressent plus au *quoi* qu'au *comment* (voir Engle & Kane, 2004). Dans le cas du modèle de Baddeley, le principal symptôme de ce relatif manque d'intérêt pour l'aspect fonctionnel est l'absence totale de spécification de la façon dont l'administrateur central opère, ce qui le fait apparaître comme un homonculus de l'aveu de l'auteur lui-même (e.g. Baddeley, 2002, 2003 ; voir page 19) : l'administrateur central est seulement livré avec une liste de missions (coordonner la performance entre deux tâches, alterner entre des stratégies de récupération en mémoire, orienter l'attention vers certain stimuli, interagir avec la mémoire à long terme), sans qu'on sache vraiment comment il les réalise.



Les travaux plus récents sur la mémoire de travail ont vu l'émergence de modèles mettant plutôt l'accent sur l'aspect fonctionnel : comment fonctionne réellement la mémoire de travail ? Autrement dit, comment se fait en pratique la maintenance de l'information en mémoire de travail ? On peut citer l'exemple du modèle de Cowan (1999), pour lequel la mémoire de travail correspond aux représentations en mémoire à long terme les plus fortement activées à un moment donné, en particulier celles vers lesquelles s'oriente le focus attentionnel. Ce modèle ne fait pas d'hypothèses particulières sur des subdivisions au sein de la mémoire de travail (il ne postule pas d'entités distinctes pour les traces mnésiques verbales et

visuo-spatiales par exemple), et s'intéresse plutôt à la façon dont les traces mnésiques déclinent, par un mécanisme de déclin ou bien d'interférence, lorsque leur activation n'est pas maintenue par le focus attentionnel.

Un autre exemple est le modèle TBRS de Barrouillet et Camos (2007), qui met l'accent sur la façon dont les traces mnésiques déclinent lorsqu'elles ne sont pas activement rafraîchies par un mécanisme attentionnel ou d'auto-répétition (Camos et al., 2009 ; Mora & Camos, 2013) ; la manipulation demandée en plus du stockage dans le cadre d'une tâche de mémoire de travail gêne ce processus de rafraîchissement, et la performance en mémoire de travail est alors fonction du temps restant disponible pour le rafraîchissement de la trace mnésique une fois la manipulation de l'information terminée. Les approches ayant pour objet de déterminer si la capacité en mémoire de travail se présente sous la forme d'emplacements (elle dispose d'un nombre limité de *slots* ; un item est maintenu s'il peut s'insérer dans un emplacement libre et oublié dans le cas contraire) ou de ressources (elle dispose d'une capacité limitée en termes de ressources, réparties en quantité variable entre les différents items) sont un autre exemple de préoccupation plutôt fonctionnelle (e.g. Ma et al., 2014).

En ce qui concerne l'interprétation de la variabilité dans la capacité en mémoire de travail, l'exemple archétypal d'un modèle fonctionnel est le modèle décrit par Engle et Kane (2004). Ce modèle, économique, se résume à deux propositions : 1) "mémoire de travail = mémoire à court terme + contrôle attentionnel", et 2) les différences inter-individuelles dans la capacité en mémoire de travail sont essentiellement liées aux processus de contrôle attentionnel (Kane & Engle, 2003). La structure interne, par exemple, du stockage à court terme n'est pas spécifiée et intéresse peu les auteurs, qui mettent plutôt l'accent sur le rôle du contrôle attentionnel dans la production de la variabilité, et sur les différentes fonctions assurées par le contrôle attentionnel (aussi appelé contrôle exécutif ou contrôle cognitif selon les études). Les variantes ultérieures du modèle (Unsworth & Engle, 2007a) conservent une ligne similaire et présentent la performance en mémoire de travail comme produit du fonctionnement du contrôle attentionnel, de la mémoire primaire et de la mémoire secondaire.

De façon générale, la plupart de ces modèles sont relativement conciliables. Le modèle d'Engle et collaborateurs ne présente pas d'opposition fondamentale avec le modèle de Baddeley par exemple : "la mémoire à court terme" (ou la mémoire primaire) peut être rapprochée des systèmes esclaves, et "le contrôle exécutif" de l'administrateur central. Tous les

modèles courants de la mémoire de travail conservent à la fois la notion de stockage à court terme, et de processus attentionnels ou exécutifs servant à manipuler l'information et/ou à rafraîchir la trace mnésique. La principale divergence philosophique entre les approches structurelles et fonctionnelles est probablement la nuance dans la façon de présenter ces concepts.

Le modèle de Baddeley présente l'administrateur central et les systèmes esclaves plutôt comme des entités réalistes : il s'agirait de systèmes distincts qui existent concrètement sur le plan cognitif, mais aussi sur le plan neurologique, une part importante de la littérature s'étant intéressée à identifier le substrat neuronal des différentes composantes du modèle (en concluant souvent qu'il n'existe pas de substrat unitaire ; Collette & Van der Linden, 2002 ; Stuss, 2011). À l'inverse, les modèles fonctionnels n'imposent pas une vision réaliste des entités qu'ils décrivent : par exemple, le modèle d'Engle et Kane (2004) s'appuie sur "les processus de contrôle exécutif" plutôt que sur "l'administrateur central". Ces processus de contrôle se définissent uniquement par le rôle fonctionnel qu'ils assurent dans la performance en mémoire de travail. *L'administrateur central* est vu comme une entité causale unitaire, statique ; *les processus de contrôle exécutif* sont vus comme probablement multiples, et comme opérant de façon dynamique.

Plus encore que pour le contrôle cognitif, l'approche fonctionnelle de la mémoire de travail encourage donc une vision constructiviste qui intègre une reconnaissance de la pluralité des mécanismes en jeu dans la détermination de la performance. Les modèles structurels s'accommodent de ce caractère multi-déterminé, puisqu'une faible performance peut s'interpréter dans le cadre du modèle de Baddeley (2000) comme résultant d'un administrateur central ou d'un système esclave inefficace ; mais les déterminants en jeu sont nécessairement moins nombreux et moins complexes dans leurs interactions que pour une vision qui prend en compte un grand nombre de processus contributeurs (ce qui autorise, par exemple, des substitutions fonctionnelles entre processus). La vision fonctionnelle est aussi probablement plus encline à tenir compte des spécificités du contexte dans lequel le système évolue, qui peuvent impacter la façon dont les différents processus se déploient.

Les travaux les plus récents commencent ainsi à présenter le maintien en mémoire de travail non pas comme le produit statique de l'opération de quelques composantes structurelles bien délimitées, mais comme le résultat de l'interaction d'un ensemble de



processus assurant une fonction particulière. Autrement dit, la mémoire de travail n'est pas vue comme une structure cognitive réaliste, mais comme la propriété émergente du fonctionnement d'une constellation de processus cognitifs qui opèrent ensemble sous la forme d'un système intégré pour maintenir et manipuler l'information (e.g. Arsalidou, 2013; Kovacs & Conway, 2016 ; Postle, 2006 ; Simmering & Perone, 2013). Le nombre d'éléments que je peux maintenir simultanément en mémoire de travail ne découle pas du fonctionnement de ma "mémoire de travail", mais du fonctionnement d'un ensemble de mécanismes qui contribuent tous à produire la performance, certains étant recrutés dans toutes les tâches de mémoire de travail et d'autres pouvant être plus spécifiques à un type de tâche ou de matériel.



Comme je l'ai discuté plus haut (voir page 14), cette vision constructiviste de la mémoire de travail comme une propriété émergente a des conséquences pratiques. Une conséquence terminologique consiste à éviter le terme "mémoire de travail" pour privilégier celui de "capacité en mémoire de travail", qui met en avant le fait que l'on étudie la capacité, la performance d'un système complet, plutôt que la propriété d'une fonction bien délimitée. Au niveau causal, la reconnaissance du caractère multi-déterminé de la mémoire de travail induit nécessairement un scepticisme quant à ses relations avec les autres construits (Gonthier, 2016) : n'importe lequel des processus contribuant à la performance en mémoire de travail peut voir son efficacité corrélée avec une autre tâche. Ce constat se marie en réalité assez bien avec la grande généralité des corrélations observées - puisque les tâches de mémoire font appel à des aspects très généraux, aussi bien mnésiques qu'attentionnels qu'ayant trait au traitement d'informations demandé par la tâche, il est logique que ces tâches corrélerent avec tout et n'importe quoi. Par ailleurs, voir la mémoire de travail comme la propriété émergente d'une constellation de processus induit nécessairement un scepticisme quant aux tentatives d'entraîner la mémoire de travail, aujourd'hui très à la mode et pourtant peu efficaces (Melby-Lervåg & Hulme, 2013) : cette inefficacité apparaît comme relativement logique, dans la mesure où l'on essaie d'entraîner une performance sans agir sur les processus causaux qui en sont à l'origine - ce qui revient à essayer d'augmenter la vitesse d'une voiture en la faisant beaucoup rouler.



Il n'existe pas encore, je pense, de modèle complet des déterminants de la performance en jeu dans une tâche de mémoire de travail. On peut toutefois proposer une petite liste des principaux déterminants identifiés ou proposés par la littérature comme contribuant



directement aux différences inter-individuelles de capacité en mémoire de travail (e.g. Gonthier, 2016) :

1) La mémoire primaire, c'est-à-dire le fait de pouvoir stocker des informations pendant une durée limitée pendant la tâche (e.g. Engle et al., 1999 ; Conway et al., 2002 ; Unsworth & Engle, 2007a).

2) La mémoire secondaire, c'est-à-dire le fait de pouvoir stocker une partie des informations en mémoire à long terme, notamment lorsque les demandes de la tâche dépassent les capacités de la mémoire primaire.

3) Le contrôle cognitif (Engle & Kane, 2004), qui peut servir de multiples rôles dans une tâche de mémoire de travail, ayant trait notamment au besoin de maintenir l'attention de façon soutenue et sélective sur les stimuli à mémoriser tout en alternant avec la manipulation de l'information.

4) Le niveau de traitement de l'information à encoder, c'est-à-dire la profondeur d'encodage des cibles - encodage visuel, phonologique, sémantique ? (Rose et al., 2012 ; Rose, 2013).

5) L'aptitude à mémoriser spécifiquement le type de matériel en jeu dans la tâche : par exemple, les aptitudes visuo-spatiales pour une tâche de mémoire de travail visuo-spatiale (Kane et al., 2004).

6) Au moins deux mécanismes de rafraîchissement des traces mnésiques : un mécanisme général de rafraîchissement attentionnel, et un mécanisme d'autorépétition subvocale spécifique au domaine verbal (Camos et al., 2009 ; Mora & Camos, 2013).

7) La stratégie utilisée dans le cadre de la tâche : autorépétition, génération de phrases, chunking, imagerie mentale, etc (Dunlosky & Kane, 2007).

8) Les compétences domaines-spécifiques en lien avec la manipulation de l'information demandée par la tâche (e.g. Süß et al., 2002 ; Kane et al., 2004), dont on peut penser que plus elles sont élevées, plus la manipulation est réalisée rapidement, laissant ainsi plus de temps disponible pour le rafraîchissement de la trace mnésique (Barrouillet & Camos, 2007).

9) La vitesse de traitement générale (Fry & Hale, 2000 ; Ackerman et al., 2002), pour les mêmes raisons, si tant est que la notion de vitesse de traitement ait un sens (Danthiir et al., 2005).

10) Une aptitude spécifique à la coordination entre plusieurs tâches, c'est-à-dire une aptitude au multitasking, permettant d'alterner entre maintien et rafraîchissement (Buehner et al., 2006).

11) La motivation à réussir la tâche (Brose et al., 2012), que l'on peut rapprocher de traits de personnalité / traits motivationnels comme le besoin de cognition (Fletcher et al., 2011).

12) Le binding, c'est-à-dire l'aptitude à intégrer chaque stimulus à mémoriser à l'emplacement correct, que ce soit son emplacement sériel ou son emplacement dans l'espace par exemple (Wilhelm et al., 2013) ; aptitude que l'on peut également aborder sous l'angle de l'aptitude à maintenir l'ordre sériel des stimuli à mémoriser (Majerus, 2013).

13) L'arousal, c'est-à-dire l'aptitude à maintenir un niveau d'activation élevé et approprié à la tâche en cours (Robison & Brewer, 2020).

14) La vitesse articulatoire, probablement impliquée dans le mécanisme d'autorépétition subvocale (Baddeley, 1986 ; Cowan et al., 1998), quoiqu'elle corrèle également avec la performance en mémoire de travail visuo-spatiale (e.g. Hamilton et al., 2003).

Cette liste n'est pas exhaustive car certains mécanismes en jeu ne sont (presque) jamais discutés par la littérature, comme par exemple la perception correcte des stimuli qui constituent la tâche, le balayage visuel des stimuli, le traitement de la séquentialité, la réalisation d'une réponse motrice ; d'autres déterminants peuvent se recouvrir partiellement, comme l'aptitude à la coordination entre les tâches et le contrôle attentionnel, ou bien la stratégie employée et la profondeur d'encodage, sans être strictement équivalents. L'étendue de cette liste suffit, à mon avis, pour établir le caractère multi-déterminé de la capacité en mémoire de travail et donc la pertinence d'une approche constructiviste plutôt que réaliste.

L'étude des aires cérébrales impliquées dans la mémoire de travail donne une image similaire (Collette & Van der Linden, 2002 ; Henson, 2001 ; Owen et al., 2005 ; Paulesu et al., 1993 ; Stuss, 2011) : on retrouve des rôles du cortex préfrontal (région dorsolatérale impliquée dans le contrôle cognitif, région ventrolatérale impliquée dans la récupération en mémoire, région rostral impliquée dans la coordination), du cortex frontal (aire de Broca et régions motrices mises en jeu par le traitement des stimuli verbaux) ; du cortex pariétal postérieur dorsal (impliqué dans le stockage et le rafraîchissement de l'information) et ventral (impliqué dans le traitement des stimuli visuo-spatiaux) ; du cortex temporal pour les tâches verbales, et

du cortex occipital pour les tâches visuelles ; et même du cervelet. Autrement dit, la mémoire de travail semble moins constituer une fonction cognitive précisément localisée qu'un cas de *neural reuse* (Anderson, 2010) : un phénomène qui émerge du fonctionnement, de la réutilisation, d'un grand nombre de réseaux de neurones non-spécifiques, recrutés de façon variable selon les demandes précises d'une tâche - là encore à l'appui d'une position constructiviste.



2. Évaluer la capacité en mémoire de travail

2.1. En tenant compte de son caractère domaine-spécifique

Un point de débat récurrent concernant la capacité en mémoire de travail est de savoir si elle est modale ou amodale : dépend-elle du type de matériel avec lequel on la mesure ? Autrement dit, existe-t-il une différence substantielle entre capacité en mémoire de travail verbale et visuo-spatiale par exemple ? De nombreux modèles théoriques ont postulé une distinction majeure entre modalités sensorielles. Le modèle de Baddeley (1986), bien sûr, qui propose une distinction structurelle entre système esclave dédié au maintien d'informations verbales (la boucle phonologique) ou visuo-spatiales (le calepin visuo-spatial) ; quoique cette distinction puisse être questionnée dans la mesure où ces deux composants sont décrits comme ayant une architecture et un fonctionnement très similaires (Logie, 1995). Mais aussi des modèles moins généraux, comme le modèle en continuum de Cornoldi et Vecchi (2003), pour lequel les processus de maintien de l'information varient fondamentalement en fonction de caractéristiques modalités-spécifiques des stimuli ; le modèle O-OER (Jones et al., 1996) dédié à la mémoire de travail verbale-auditive ; ou encore les modèles *object files* (Kahneman & Treisman, 1984) et *object tokens* (Marr, 1976) dédiés à la mémoire de travail visuelle.

D'un point de vue empirique, plusieurs arguments parlent en faveur de l'existence de systèmes de stockage distincts (Jonides et al., 1996). Le plus discuté par la littérature repose sur les paradigmes de double tâche, qui montrent une interférence plus importante lorsque deux tâches concurrentes impliquent du matériel différent, notamment verbal et visuo-spatial (Kroll et al. 1970), mais aussi visuel et spatial (Klauer & Zhao, 2004). L'existence de doubles dissociations en neuropsychologie, avec des déficits en mémoire de travail visuelle mais pas verbale et inversement (e.g. Jonides et al., 1996 ; Wang & Bellugi, 1994), et l'existence d'activités cérébrales distinctes en fonction du type de matériel (e.g. Mecklinger & Müller, 1996 ; Smith, Jonides, & Koeppel, 1996) sont d'autres bons indices que mémoire de travail verbale et visuo-spatiale ne recouvrent pas tout-à-fait les mêmes processus.

D'un autre côté, les processus de contrôle cognitif vus comme critiques pour la capacité en mémoire de travail sont conçus comme étant domaine-généraux plutôt que spécifiques à une modalité sensorielle (e.g. Baddeley, 2003 ; Kane et al., 2004) ; les réseaux de neurones mis en jeu par les tâches de mémoire de travail qui relèvent de modalités différentes se

chevauchent largement (Chein et al., 2011 ; Majerus et al., 2010). Une littérature abondante indique que les scores à des tâches de mémoire de travail qui appartiennent à des modalités sensorielles différentes sont fortement corrélés, et tendent à saturer un même facteur domaine-général plutôt que des facteurs différents selon le type de matériel (e.g. Ackerman et al., 2002 ; Conway et al., 2002 ; Engle et al., 1999 ; Kane et al., 2004 ; Oberauer et al., 2003 ; Süß et al., 2002). Par ailleurs, les tâches de mémoire de travail verbales et visuo-spatiales tendent à montrer les mêmes corrélations avec d'autres tâches (Conway et al., 2002 ; Engle et al., 1999 ; Kane et al., 2004).



Ce tableau se marie bien avec l'approche constructiviste : la capacité en mémoire de travail émerge du fonctionnement de processus transversaux impliqués quelle que soit la tâche, et de processus spécifiques au type de matériel. Les travaux les plus récents tendent à proposer l'existence d'une mémoire de travail basée sur des processus, notamment attentionnels, essentiellement généraux, éventuellement adossés à des mécanismes spécifiques dédiés aux stimuli verbaux (Camos et al., 2009 ; Morey, 2018).

Si l'on suit la proposition d'Engle et Kane (2004) selon laquelle ce sont surtout les processus de contrôle cognitif domaine-généraux qui créent la valeur prédictive de la mémoire de travail (si la capacité en mémoire de travail corrèle avec d'autres tâches cognitives de haut niveau, c'est surtout à cause du rôle du contrôle cognitif), alors la façon la plus appropriée d'évaluer la capacité en mémoire de travail consiste à combiner plusieurs tâches qui s'appuient sur des modalités sensorielles différentes. Cette stratégie a régulièrement été utilisée dans le cadre des études en équations structurales (e.g. Engle et al., 1999 ; Conway et al., 2002 ; Kane et al., 2004), sans vraiment être formalisée en tant que telle. C'est dans ce cadre que j'ai développé le premier empan complexe composite de mémoire de travail, le Composite Complex Span (CCS ; Gonthier, Thomassin, & Roulin, 2016), qui intègre trois subtests : un empan complexe de lecture, un empan complexe de symétrie, et un empan complexe d'opérations.

Les empan complexes sont la méthodologie la plus utilisée pour l'étude des différences inter-individuelles de capacité en mémoire de travail (Conway et al., 2005 ; Redick et al., 2012). Ces tâches alternent entre présentation d'un stimulus à mémoriser et présentation d'une tâche concurrente de traitement de l'information (décider si une phrase est correcte, si une image est symétrique, si une opération mathématique est correcte, respectivement, pour les empan de

lecture, de symétrie et d'opérations) ; elles répondent ainsi directement à la définition de la mémoire de travail comme l'association entre mémorisation et manipulation de l'information. Pour créer cette tâche, j'ai combiné les trois empan complexes les plus utilisés, qui avaient fait l'objet d'une validation en anglais (Unsworth et al., 2005), en m'appuyant sur les stimuli originaux : il s'agit d'un empan de phrases (mémoriser des stimuli tout en déterminant si des phrases sont correctes; Daneman & Carpenter, 1980), d'un empan de symétrie (mémoriser des stimuli visuo-spatiaux tout en déterminant si des images présentent une symétrie verticale ; Kane et al., 2004), et d'un empan d'opérations (Turner & Engle, 1989).

Le CCS a quelques originalités. En plus d'être la première batterie composite d'empan complexes, il s'agit de la première validation en langue française d'empan complexes qui respectent la même structure que les versions couramment utilisées en anglais (les tâches d'empan de lecture de Desmette et al., 1995, ou encore Delaloye et al., 2008, utilisent des formats de phrases assez différents, tandis que les versions de l'empan d'opérations de Fournet et al., 2012, font mémoriser des stimuli plus complexes tels que des mots). Le score sur le CCS, supposé refléter la mémoire de travail domaine-générale, est calculé comme la moyenne des scores de rappel sur chacun des trois subtests (après transformation en scores z). Pour optimiser la sensibilité de la tâche, le scoring sur chaque subtest se fait comme la somme totale du nombre de stimuli rappelés dans la bonne position séquentielle à travers l'ensemble des essais, conformément aux recommandations (Conway et al., 2005), plutôt qu'en cotant les essais comme ratés ou réussis (en tout-ou-rien).

Le CCS ne permet pas de calculer un score par subtest. Contrairement aux empan complexes anglophones utilisés pour les études en équations structurales (Unsworth et al., 2005), prévus pour être des tâches indépendantes, le CCS a été conçu directement comme l'association de trois subtests, ce qui a permis de raccourcir considérablement chacun des trois subtests par rapport à leurs versions originales : le nombre d'essais par subtest est limité et la fidélité du score par subtest est également, ce qui n'est pas un problème puisque seul le score global est d'intérêt. Conformément à l'idée qu'une tâche doit présenter plus d'essais de difficulté intermédiaire (pour discriminer entre les nombreux participants d'aptitude moyenne) que d'essais faciles ou difficiles (pour discriminer entre les quelques participants d'aptitude faible ou élevée), chaque subtest ne présente qu'un seul essai pour les niveaux de difficulté extrême, ce qui permet de raccourcir encore le temps total de passation. Ce temps s'élève à

environ 25-30 minutes - nettement moins que pour l'association de trois empan complexes présentés dans leur version intégrale (dont la passation est plutôt de l'ordre de 45 minutes).

La validation du CCS sur un peu plus de 1000 participants (étudiants en psychologie) a montré une excellente consistance interne (alpha de Cronbach entre .70 et .76 pour chaque subtest et omega total de McDonald de .86 pour le score total) et une bonne fidélité test-retest (.77 pour le score total), ainsi qu'une très bonne validité ($r = .39$ avec les matrices de Raven, très proche de la corrélation de .36 reportée dans les méta-analyses ; Ackerman et al., 2005). La collecte d'un échantillon de validation important m'a également permis de définir des normes, qui permettent de coter les résultats de nouveaux participants en référence à l'échantillon de validation initial. C'est notamment utile pour standardiser le score sur chaque subtest (qui est z-scoré par rapport à l'échantillon global de validation) avant de calculer le score total, mais aussi pour déterminer des critères d'exclusion liés à la précision sur les tâches concurrentes de manipulation de l'information : les étudiants français apparaissent en effet nettement plus faibles que les étudiants américains, notamment en ce qui concerne la précision sur les vérifications d'opérations mathématiques (cinquième percentile = 57.5% de réponses correctes, contre environ 85% pour la version originale de l'empan d'opérations).

Le CCS n'est pas exempt de limitations ; je relèverais trois défauts principaux à cette tâche. Le premier est qu'en combinant les trois empan complexes les plus utilisés, j'ai aussi déséquilibré la représentation des modalités sensorielles : empan de lecture et empan d'opérations demandent de mémoriser du matériel verbal (des chiffres et des consonnes, respectivement), tandis que l'empan de symétrie demande de mémoriser du matériel visuo-spatial (des positions spatiales dans une grille). Il y a donc deux tâches verbales pour une tâche visuo-spatiale, et le score composite global est ainsi un peu plus verbal que visuo-spatial. C'est pour cette raison que le score intègre à égalité les trois subtests, sans les pondérer par leurs saturations respectives sur un facteur général de mémoire de travail (ce qui aboutirait logiquement à sur-pondérer les deux tâches verbales). Un second défaut est la non-prise en compte du caractère tâche-spécifique de la capacité en mémoire de travail : les trois subtests utilisent le même paradigme et intègrent donc certainement une part de variance méthodologique, spécifique à l'empan complexe (j'y reviendrai un peu plus loin, page 102). Enfin, la tâche ne s'adresse qu'à des adultes de bon niveau intellectuel. Le niveau de difficulté en est élevé (jusqu'à 8 stimuli à mémoriser pour l'empan de lecture), surtout pour les tâches de



traitement : les phrases à vérifier sont complexes (*Le ciel prit de délicates teintes pâles et le soleil glissa lentement vers l'horizon.*), les opérations également ($[9/3] + 7 = 10$: correct ?). Le CCS est donc adapté pour l'évaluation d'étudiants universitaires mais montre rapidement ses limites pour d'autres populations. Ce problème fait l'objet de la section suivante.

2.2. En tenant compte de la longueur de la tâche et du niveau d'aptitude

La recherche sur la variabilité impose régulièrement de comparer des groupes de niveaux d'aptitude très différents : c'est notamment le cas dans le cadre de la recherche développementale, mais aussi dans la recherche sur la déficience intellectuelle ou le haut potentiel. Dans le cas des tâches de mémoire de travail, cela pose une contrainte importante ; pour éviter l'apparition d'effets plancher ou plafond, la tâche doit être de difficulté adaptée pour des participants d'aptitude très faible ou très élevée. Idéalement, il faudrait donc beaucoup d'essais à tous les niveaux de difficulté (trois essais avec un stimulus à mémoriser ; trois essais avec deux stimuli à mémoriser ; etc). Cela exclut d'emblée l'usage de tests comme le CCS, calibrés pour être d'un niveau de difficulté intermédiaire *pour un échantillon d'un certain niveau d'aptitude*.

La solution classique à ce problème est la procédure ascendante, la plus utilisée en neuropsychologie notamment : on commence par les essais les plus simples, et on augmente progressivement la longueur des séquences à mémoriser. La tâche est interrompue lorsque le participant échoue une certaine proportion des essais pour un niveau de difficulté donné. (La procédure ascendante va généralement de pair avec la méthode de scoring en tout-ou-rien, déconseillée puisque peu sensible ; mais il existe des variantes où l'on interrompt la tâche lorsque le participant rappelle une proportion trop faible des stimuli.) Cette procédure permet bel et bien de présenter beaucoup d'essais à tous les niveaux de difficulté, mais elle a toutefois une conséquence perverse : elle génère une très forte variabilité dans la longueur de la tâche. Si l'on prend l'exemple d'un subtest de la batterie AWMA (Alloway, 2007) : un participant réalise entre trois et six essais par niveau de difficulté (on passe au niveau supérieur lorsque trois séquences ont été rappelés correctement), et les niveaux vont de 1 (un seul stimulus à mémoriser) à 7. En conséquence, un participant devra réaliser entre 3 essais et 42 essais selon son niveau d'aptitude. En termes de stimuli, un participant devra mémoriser entre 3 stimuli et 168 stimuli pour ce subtest. Dans la mesure où l'AWMA comporte 12 subtests, on mesure l'énorme différence que l'on peut obtenir dans la longueur de la tâche selon le niveau d'aptitude du participant évalué.

D'un point de vue constructiviste, cette variation dans la longueur de la tâche est très problématique car elle peut agir de façon différentielle sur les différents processus qui contribuent à la détermination de la performance ; et le problème est double puisque cette

variation est confondue avec le niveau d'aptitude. Au-delà du simple problème organisationnel posé par l'imprévisibilité du temps de passation, voici quelques problèmes significatifs : 1) La tâche génère des différences de fatigue importantes d'un participant à l'autre, qui pourront de plus affecter la performance sur des tâches subséquentes et créer ou masquer des relations entre construits (un participant d'aptitude élevée en mémoire de travail étant plus fatigué lorsqu'il passe à la seconde tâche). 2) La tâche génère une lassitude différente d'un participant à l'autre, qui peut elle-même induire un désengagement : on peut s'attendre à ce qu'un participant à haut potentiel se lasse de la tâche bien avant d'avoir mémorisé 168 stimuli x 12 subtests, et on ne serait pas surpris de le voir échouer volontairement pour interrompre la passation (Vandierendonck et al., 1998). 3) La tâche génère une interférence proactive différente d'un participant à l'autre : un participant d'aptitude élevée accumulera beaucoup plus d'interférence proactive puisqu'il devra mémoriser un nombre de stimuli total bien plus élevé. Or, l'interférence proactive modère la validité des tâches de mémoire de travail (Lustig et al., 2001), et la résistance à l'interférence varie en fonction de l'âge et du niveau d'aptitude (Loosli et al., 2014 ; Kane & Engle, 2000).



Au-delà de la longueur de la tâche, la procédure ascendante avec interruption est à l'origine de plusieurs problèmes : 4) Une baisse passagère d'attention ou d'effort investi, une fluctuation attentionnelle, peuvent être à l'origine d'un échec qui mène à interrompre trop tôt la passation : le score total peut être radicalement altéré par un seul essai. 5) Le caractère prévisible de la longueur de chaque essai permet la mise en place de stratégies d'anticipation qui tiennent compte du nombre de stimuli à mémoriser, ce que l'on évite habituellement par une pseudo-randomisation des niveaux de difficulté (Conway et al., 2005) ; 6) La passation est interrompue lorsque le participant échoue, ce qui veut dire que la tâche se termine systématiquement sur un échec, ce qui peut induire un découragement, notamment en contexte clinique.



Enfin, d'un point de vue psychométrique, 7) Le nombre d'essais variable d'un participant à l'autre peut augmenter artificiellement la variance d'erreur des scores (Woods et al., 2011), et implique d'ailleurs que l'erreur de mesure diffère largement d'un participant à l'autre, problème qui existe dans tous les tests mais particulièrement renforcé ici ; 8) Au final, un participant donné ne complètera que quelques essais réellement utiles, puisque seuls quelques essais sont réellement appropriés à son niveau d'aptitude (Weiss, 1974), la plupart étant trop

durs ou trop faciles. Ces nombreux problèmes limitent fortement l'intérêt de la procédure ascendante pour l'évaluation de la variabilité. La solution de niveler le temps de passation par le haut en faisant réaliser tous les essais à tous les participants ne serait guère plus satisfaisante : la tâche sera toujours très longue, la plupart des essais seront toujours inutiles, et en plus les participants d'aptitude faible seront confrontés à un très grand nombre d'échecs répétitifs.

La solution que j'ai proposée a consisté à développer une variante adaptative (Lord, 1968 ; Weiss, 1974) des tâches de mémoire de travail, sous la forme de l'Adaptative Composite ComplEx Span (ACCES ; Gonthier, Aubry, & Bourdin, 2017). La tâche elle-même reprend la structure du CCS en combinant un empan de lecture, un empan de symétrie et un empan d'opérations, de façon à obtenir un score composite de mémoire de travail. Les tâches de manipulation de l'information demandées par chaque subtest ont toutes été rendues très simples (phrases et opérations élémentaires, ajout d'un guide visuel pour juger la symétrie des images sur l'empan de symétrie), en partant du principe que la difficulté de la manipulation n'a pas d'importance tant qu'elle empêche le rafraîchissement actif des traces mnésiques (Lépine et al., 2005 ; Turner & Engle, 1989).

L'originalité de l'ACCES réside dans son caractère adaptatif, dont le principe est simple : on commence la tâche à un niveau de difficulté intermédiaire, et on ajuste la difficulté à chaque nouvel essai en fonction de la performance du participant. Cette idée est représentée dans la Figure 6. (La mise en œuvre de cette procédure est rendue très simple par le fait que les items des tâches de mémoire de travail sont bien résumés par des modèles de réponse à l'item à un seul paramètre de difficulté, lui-même bien approximé par le nombre d'éléments à mémoriser ; Vock & Holling, 2008.)

Cette procédure présente de nombreux avantages. L'ajustement dynamique de la difficulté à chaque essai permet à la tâche de converger rapidement vers un niveau de difficulté individualisé, approprié à chaque participant : un participant d'aptitude faible oscillera autour d'essais faciles tandis qu'un participant d'aptitude élevée réalisera surtout des essais difficiles. D'un point de vue psychométrique, la plupart des essais pour un participant donné seront donc diagnostiques de son niveau d'aptitude ; d'un point de vue motivationnel, la tâche n'est jamais désengageante puisqu'elle est toujours adaptée au niveau du participant. Dans le cadre de la validation de la tâche, nous l'avons fait passer à des enfants du CE2 à la 5ème, et nous avons évalué la difficulté perçue à l'aide d'un questionnaire en auto-report ; la difficulté auto-estimée

s'est avérée constante à travers les âges, et les enfants ont donné de bons retours qualitatifs. Un avantage majeur de cette procédure est qu'elle permet de faire passer exactement la même tâche standardisée à des participants de niveaux d'aptitude très différents : j'ai fait passer cette tâche à des enfants de CE2, à des étudiants en licence de psychologie, et à des personnes âgées jusqu'à 80 ans, sans jamais observer d'effet plancher ou d'effet plafond.

L'ACCES a fait l'objet d'une validation psychométrique dans un échantillon de 268 enfants âgés de 8 à 13 ans ; cette procédure pose toutefois quelques problèmes de psychométrie, puisque chaque participant passe des essais légèrement différents. Cela empêche donc l'usage, par exemple, d'analyses factorielles, ou l'évaluation de la consistance interne. La tâche a tout de même montré une bonne sensibilité, une fidélité test-retest a minima (.70), et une bonne corrélation avec les matrices de Raven ($r = .34$). Elle est librement utilisable et dispose même de normes pour les enfants, issues de cet échantillon de validation (Gonthier et al., 2016) ; mais je vois moins l'ACCES comme la version définitive d'une tâche francophone que comme une proof-of-concept de la faisabilité d'un paradigme adaptatif. La procédure adaptative est très simple à mettre en œuvre dans ce contexte et pourrait facilement se décliner à d'autres tâches.

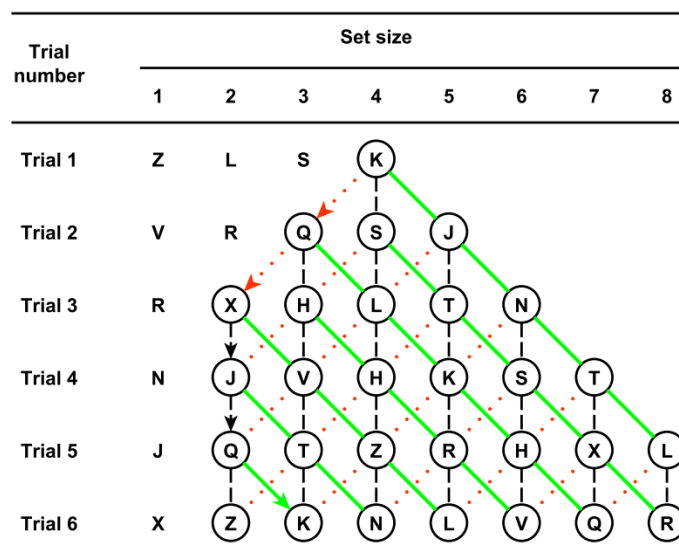


Figure 6. Trajectoires possibles pour un subtest de l'ACCES. Le premier essai demande au participant de mémoriser la séquence ZLSK ; selon le degré de réussite, l'essai suivant demandera de mémoriser la séquence VRQ, VRQS, ou bien VRQSJ. Tiré de Gonthier, C., Aubry, A., & Bourdin, B. (2017). Measuring working memory capacity in children using adaptive tasks: Example validation of an adaptive complex span. *Behavior Research Methods*, 50(3), 910-912.

3. Mécanismes de la performance en mémoire de travail

3.1. Étapes du processus mnésique

Les principales étapes du processus mnésique sont bien connues : encodage du stimulus à mémoriser, stockage et maintien en mémoire, et récupération en mémoire. Ce schéma de base se décline de façon équivalente en mémoire à long terme, en mémoire à court terme et en mémoire de travail. Les quelques complexités supplémentaires qui s'ajoutent à ce schéma dans le cadre de tâches de mémoire de travail (la phase de maintien a lieu simultanément à la phase d'encodage lorsque les stimuli sont présentés séquentiellement ; Gonthier, 2020 ; la phase de maintien en mémoire doit être réalisée en concurrence avec une manipulation de l'information ; Barrouillet & Camos, 2007) ne changent pas ce processus de base.

Le développement de méthodes de neuroimagerie a certes généré une abondance de recherches concernant la *structure* des régions cérébrales impliquées dans les tâches de mémoire de travail ; mais une autre série de travaux, moins connue, s'est intéressée au *fonctionnement* du processus mnésique. Ces travaux sont essentiellement issus de l'électro-encéphalographie, dont l'intérêt dans ce contexte est sa très forte résolution temporelle (extraction de l'activité neuronale avec une précision de quelques millisecondes). Cette résolution temporelle élevée permet d'isoler différentes étapes fonctionnelles dans le cadre d'une tâche de mémorisation, en identifiant des potentiels évoqués qui émergent de façon distincte à différents moments du processus. Il n'est pas toujours simple de mettre en relation un potentiel évoqué avec une activité cognitive précise, mais on peut obtenir des informations à ce niveau en manipulant différents aspects de la tâche ou en étudiant, par exemple, la corrélation entre un potentiel évoqué et la performance de rappel dans un essai donné.



Cette approche a permis de mettre en lumière une complexité du processus mnésique un peu plus importante que les trois étapes traditionnellement retenues. J'ai réalisé une synthèse de ces apports théoriques (Gonthier & Hot, 2013), dont le résumé apparaît dans la Figure 7. Ces résultats ont été obtenus avec un mélange de tâches de mémoire à court ou à long terme (la distinction n'est pas toujours claire au vu des paradigmes utilisés en électro-encéphalographie) et se généralisent probablement à la mémoire de travail. On relève au moins trois étapes supplémentaires dans le processus de mémorisation. D'une part, la phase d'encodage est précédée d'une étape de pré-encodage, qui commence environ 1 seconde

avant la présentation du stimulus à mémoriser. L'amplitude des potentiels évoqués pendant cette étape de pré-encodage d'un stimulus est directement en lien avec la qualité du rappel subséquent ; il s'agit donc d'une étape qui joue un vrai rôle fonctionnel (e.g. Otten et al., 2006). Trois interprétations coexistent dans la littérature : cette étape de pré-encodage peut constituer une forme d'encodage du contexte de présentation du stimulus, qui faciliterait le processus d'écphorie synergétique ; elle pourrait correspondre à la mise en jeu préparatoire de processus, notamment sémantiques, qui favorisent un encodage profond du stimulus ; ou elle pourrait refléter la mobilisation attentionnelle du participant qui se prépare à encoder le stimulus. Cette dernière hypothèse peut se reformuler en termes de motivation ou d'arousal, et est encouragée par le constat que cette activité pré-encodage est modérée par des variables motivationnelles (Gruber & Otten, 2010).

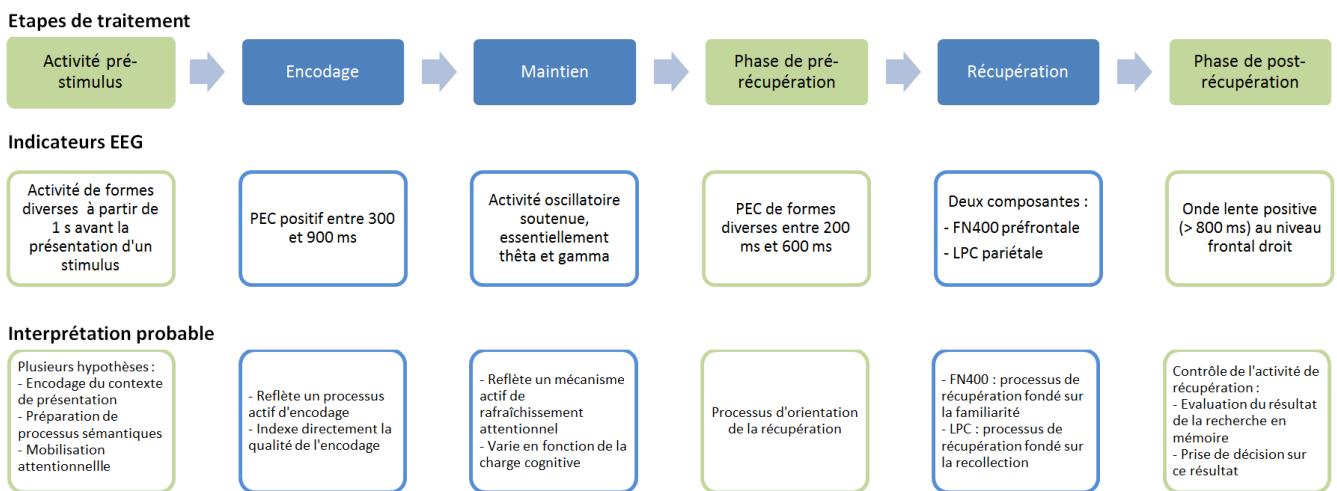


Figure 7. Étapes de traitement de l'information dans le cadre d'une tâche de mémoire de travail : étapes classiques (en bleu) et étapes supplémentaires suggérées par l'analyse de l'activité électro-encéphalographique (en vert). Tiré de Gonthier, C., & Hot, P. (2013). Apports de l'électro-encéphalographie à la compréhension de la mémoire. *Revue de neuropsychologie*, 5(4), 243-254.

Une seconde étape supplémentaire se situe juste avant la phase de récupération : on observe des potentiels évoqués spécifiques dans les quelques centaines de millisecondes qui précèdent la récupération d'un stimulus en mémoire. La forme de ces potentiels évoqués dépend du type de tâche (par exemple identification d'un item parmi des distracteurs ou rappel du contexte de présentation). Cette activité apparaît même lorsque l'on demande au participant d'identifier un distracteur de même type que les stimuli à mémoriser, et ne peut

donc pas correspondre à une récupération en mémoire ; elle reflète plutôt un processus préparatoire à la récupération lié au type de récupération demandée (Rugg & Wilding, 2000). On la qualifie "d'orientation de la récupération" (Rugg & Wilding, 2000 ; Jaeger & de Parente, 2008) : l'idée est qu'une tâche de récupération particulière en mémoire demande une orientation vers ce type de traitement en particulier. L'activité d'orientation de la récupération reflèterait donc une sorte d'aiguillage mental, une mise en branle du type de processus demandé par la tâche.

Enfin, la phase de récupération est suivie d'une phase post-récupération qui dure environ 1 seconde (Friedman & Johnson, 2000) et peut même perdurer après la réponse du participant. On l'attribue classiquement à un processus de monitoring de la récupération et de décision comportementale, d'ordre métacognitif : ma récupération en mémoire est-elle fiable ? dois-je vraiment donner cette réponse (Jaeger & Parente, 2008) ? Cette activité post-récupération n'est d'ailleurs pas spécifique à la mémoire et s'observe également dans d'autres types de tâches qui demandent un jugement, une prise de décision (Hayama et al., 2008).

En plus de mettre en relief la complexité fonctionnelle du processus mnésique, les études en électro-encéphalographie ont aussi largement montré son caractère modulaire : on retrouve systématiquement une combinaison d'activités corticales spécifiques au type de tâche en cours et d'activités générales indépendantes de la tâche (Gonthier & Hot, 2013). C'est notamment vrai pour différents types de tâches et de stimuli - verbal, visuo-spatial ou encore haptique - qui génèrent d'abord des activités différentes dans les aires sensorielles primaires à l'encodage, mais semblent ensuite recodés sous forme de représentations mnésiques plus ou moins amodales (Ruchkin et al., 1997). La phase de maintien semble montrer une activité relativement indépendante du type de stimulus (e.g. Gevins et al., 1997), avec une activité supplémentaire dans le cas de stimuli verbaux (Hwang et al., 2005), probablement en lien avec le mécanisme d'autorépétition subvocale spécifique à ces stimuli, mais aussi une activité spécifique au maintien de l'ordre sériel (Roberts et al., 2013). On aboutit à des conclusions similaires pour la phase de récupération (McEvoy et al., 1998). De la même façon, tâches de mémoire sémantique et de mémoire épisodique (Johnson et al., 1998), tâches de rappel et de reconnaissance (Allan & Rugg, 1997), révèlent une association d'activités corticales spécifiques et d'activités communes qui parlent fortement en faveur d'une vision constructiviste de la mémoire comme émergeant de processus plus ou moins partagés par différentes tâches.



3.2. Le rôle du contrôle cognitif

Tous les modèles théoriques de la mémoire de travail s'accordent à donner un rôle prépondérant au contrôle cognitif (généralement appelé contrôle attentionnel ou contrôle exécutif dans ce contexte). Le problème est de déterminer quels sont exactement les rôles fonctionnels du contrôle cognitif dans une tâche de mémoire de travail. Comme je l'ai évoqué dans la section précédente (page 68), l'une des critiques majeures du modèle de Baddeley est le statut homonculaire de l'administrateur central, qui "gère la performance" sans que l'on sache vraiment comment ; mais même ses fonctions dans la tâche ne sont pas clairement identifiées, en particulier lorsqu'il s'agit de déterminer laquelle ou lesquelles de ces fonctions contribuent aux différences inter-individuelles de capacité en mémoire de travail.



La liste de fonctions proposées par la littérature est vertigineuse. Dans une tâche de mémoire de travail, le contrôle cognitif permettrait de : porter son attention sur les cibles à encoder (Engle & Kane, 2004), augmenter l'activation des traces mnésiques à l'encodage (Barrouillet et al., 2009), générer et utiliser des indices de récupération en mémoire secondaire (Unsworth & Engle, 2007a), implémenter des stratégies d'encodage efficaces (Ang & Lee, 2010), résister à l'interférence proactive (Kane & Engle, 2000), gérer la coordination de l'alternance entre maintien et manipulation de l'information (Barrouillet & Camos, 2007), résister à l'interférence générée par des distracteurs (Engle & Kane, 2004), réaliser rapidement la manipulation de l'information concurrente à la maintenance (Barrouillet, Lépine, & Camos, 2008), ou encore éviter le mind-wandering (Adam et al., 2015).

Le principal problème de cette liste est que la plupart de ces hypothèses sont extrêmement difficiles à tester. Non pas parce qu'il est difficile de collecter une mesure de l'efficacité sur n'importe laquelle de ces fonctions, et de la corrélérer avec la performance en mémoire de travail : cette démarche peut poser quelques problèmes expérimentaux mais reste relativement triviale. Le principal obstacle est le caractère multi-déterminé de la mémoire de travail. Dans une perspective constructiviste, la capacité en mémoire de travail émerge du fonctionnement d'un grand nombre de processus ; cela signifie que le score est multi-déterminé (tous les déterminants contribuent à la performance) et multi-réalisable (il y a plusieurs façons d'obtenir un même score ; voir page 14), mais aussi que les corrélations qui impliquent ce score sont intrinsèquement suspectes car leur origine causale est impossible à

déterminer. C'est d'autant plus vrai que le contrôle cognitif peut lui-même être vu comme une propriété émergente plutôt que comme une entité réaliste.

On peut illustrer facilement cet argument par le fait que la performance en mémoire de travail corrèle bien avec la vitesse de traitement (Ackerman et al., 2002) et avec la motivation (Brose et al., 2012), qui peuvent bénéficier à n'importe quelle tâche complexe. On peut donc tout aussi bien imaginer qu'une motivation et une vitesse de traitement élevées soient utiles à la fois sur les tâches de mémoire de travail et sur les tâches qui évaluent les fonctions suscitées du contrôle cognitif, créant ainsi une corrélation factice entre les deux. De ce fait, comment être sûr que la corrélation entre capacité en mémoire de travail et l'une des fonctions du contrôle cognitif est bel et bien due au fait que le contrôle cognitif remplit cette fonction dans une tâche de mémoire de travail ? Pour toute la liste de fonctions potentielles du contrôle cognitif évoquées ci-dessus, on peut s'attendre à observer une corrélation avec la capacité en mémoire de travail : comme je l'ai évoqué plus haut, un très grand nombre de performances dans les domaines les plus divers corréleront avec la mémoire de travail (page 66), sans que cela fournisse beaucoup d'informations sur les mécanismes en jeu. Il est même très difficile de trouver des exemples de performances qui corréleront *négativement* avec la capacité en mémoire de travail (les paradigmes d'oubli dirigé sont l'une des rares exceptions ; Delaney & Sahakyan, 2007 ; ainsi que le phénomène de choking under pressure ; Beilock & DeCaro, 2007 ; pour un autre exemple, voir Goodhew, 2020).



Dans le cadre de ma thèse de doctorat (Gonthier, 2014), j'ai cherché à contourner ce problème en abordant le contrôle cognitif sous l'angle du modèle DMC présenté dans le chapitre précédent. Ce projet a consisté à tester le lien entre capacité en mémoire de travail et contrôle cognitif : comme je l'ai évoqué plus haut (page 25), l'implémentation du contrôle de façon préparatoire demande théoriquement le maintien actif d'informations contextuelles en mémoire de travail dans le délai qui précède l'apparition du stimulus à contrôler, et il serait donc logique qu'une forte capacité en mémoire de travail aille de pair avec une plus grande tendance à utiliser le contrôle proactif. L'intérêt de cette approche est double : d'une part, le modèle DMC est un modèle défini au niveau individuel, sous l'angle des mécanismes en jeu (et non au niveau inter-individuel comme "la performance d'inhibition" par exemple), ce qui permet des prédictions comportementales fines. D'autre part, le modèle DMC va de pair avec la tâche AX-CPT qui doit montrer une *baisse* de performance sur les essais AY lorsqu'un

participant utilise le contrôle proactif ; on devrait donc pouvoir obtenir un test véritablement diagnostique du lien entre capacité en mémoire de travail et usage d'un contrôle proactif en observant, en contrecarrant la corrélation positive observée entre mémoire de travail et performance dans tous les domaines.

Ce projet a été un échec massif (Gonthier, 2014 ; Gonthier, Baciú, & Roulin, 2018 ; Gonthier, Thomassin, & Roulin, 2012 ; Gonthier, Thomassin, Baciú, & Roulin, 2013 ; Gonthier, Thomassin, Baciú, & Roulin, 2014). À travers 15 expériences impliquant environ 1500 participants, je n'ai jamais réussi à mettre en évidence le pattern attendu, les participants à forte capacité en mémoire de travail restant systématiquement meilleurs que les participants à faible capacité, et la capacité en mémoire de travail ne montrant aucune corrélation avec la performance AY ou avec le PBI sur l'AX-CPT. On peut imaginer que cet échec soit dû en partie à la faible fidélité des mesures, en partie à la très faible variabilité du contrôle cognitif chez de jeunes adultes sains et étudiant à l'université (Cooper, Gonthier, et al., 2017). Plusieurs chercheurs se sont intéressés à cette hypothèse par la suite (Redick & Engle, 2011 ; Redick, 2014 ; Richmond et al., 2015), en arrivant à la conclusion opposée et en avançant l'existence d'un lien entre mémoire de travail et contrôle cognitif chez le jeune adulte ; mais cette conclusion s'appuie sur le constat que les participants à forte capacité en mémoire de travail sont meilleurs sur les essais BX et équivalents sur les essais AY, ce qui donne effectivement lieu à un "score de contrôle proactif plus élevé" mais n'est pas vraiment diagnostique, pour les raisons évoquées ci-dessus (voir aussi pages 53-54).



Dans le cadre d'un projet un peu différent (Gonthier & Gavornikova-Baligand, en révision), j'ai cherché à tester une hypothèse inspirée du modèle TBRS (voir page 68). Dans la perspective du modèle TBRS, le rôle central du contrôle cognitif serait d'assurer le rafraîchissement des traces mnésiques et de coordonner l'alternance entre rafraîchissement des traces mnésiques et traitement concurrent de l'information. On peut donc représenter la performance en mémoire de travail comme une fonction de la charge cognitive, telle qu'indexée par le temps disponible pour le rafraîchissement des traces mnésiques (c'est-à-dire par le ratio entre le nombre d'opérations à effectuer dans le cadre de la tâche concurrente de traitement de l'information et le temps disponible pour le faire ; voir la Figure 8A pour un exemple). Cette proposition constitue un modèle possible du rôle fonctionnel du contrôle cognitif dans une tâche de mémoire de travail, dont découle une hypothèse possible sur le rôle

du contrôle cognitif dans la variabilité : les différences de capacité en mémoire de travail pourraient être dues à une variabilité dans l'aptitude à alterner entre rafraîchissement et stockage, et/ou dans l'aptitude à rafraîchir l'information.

L'une des expériences réalisées dans le cadre du modèle TBRS (Barrouillet et al., 2009) propose un moyen possible de tester cette hypothèse : il a en effet été observé que des enfants d'âges différents diffèrent dans leur aptitude à réaliser ce rafraîchissement, cette différence se manifestant par une pente plus faible de la capacité en mémoire de travail en fonction de la charge cognitive (voir la Figure 8B). Cette différence a été interprétée comme étant due à une augmentation développementale dans l'efficacité du contrôle cognitif dans la mise en œuvre du rafraîchissement en concurrence avec la tâche interférente (quoique d'autres interprétations soient possibles, notamment une différence de vitesse de traitement ; Gaillard et al., 2011). Ainsi, la pente serait plus faible pour les enfants les plus jeunes parce qu'ils bénéficient peu d'une diminution de la charge cognitive (ou de façon équivalente, parce qu'ils sont peu gênés par l'impossibilité de rafraîchir la trace mnésique lorsque la charge cognitive augmente).

Si cette interprétation s'applique plus généralement à la variabilité, notamment chez l'adulte, on devrait observer un lien entre efficacité du rafraîchissement et différences inter-individuelles de capacité en mémoire de travail : les personnes dotées d'une capacité en mémoire de travail plus faible devraient présenter une pente plus faible pour l'effet de la charge cognitive, analogue au groupe d'enfant le plus jeune sur la Figure 8B. C'est cette hypothèse que j'ai cherché à tester en manipulant directement la charge cognitive et en vérifiant l'effet de cette manipulation sur le lien entre capacité en mémoire de travail et performance (Gonthier & Gavornikova-Baligand, en révision). Le résultat de cette expérience a été nul : les participants dotés d'une forte capacité en mémoire de travail présentent une performance uniformément supérieure, et la pente de la droite qui reflète l'efficacité de leur mécanisme de rafraîchissement est strictement équivalente à la pente de la droite pour les participants dont la capacité en mémoire de travail est la plus faible.

Je n'ai pu obtenir l'interaction recherchée entre charge cognitive et capacité en mémoire de travail qu'en utilisant une méthode de cotation en tout-ou-rien des essais, comme l'article d'origine (Barrouillet et al., 2009) ; mais l'interaction reflète dans ce cas un manque de sensibilité pour les participants à faible capacité en mémoire de travail, dont la performance



montre un effet plancher à travers les différentes conditions, ce qui crée l'illusion d'un mécanisme de rafraîchissement inefficace. Dans le cadre de la même étude, nous avons obtenu un résultat similaire avec des enfants de 8 et 14 ans : les pentes des droites qui reflètent l'efficacité du rafraîchissement sont parallèles pour les deux groupes d'âges lorsque la charge cognitive est relativement faible et que l'on cote la proportion de stimuli correctement rappelés par essai ; une interaction apparaît uniquement pour la condition de charge cognitive la plus élevée et/ou lorsque l'on cote les essais en tout-ou-rien, et reflète dans ce cas un effet plancher pour le groupe le plus jeune.

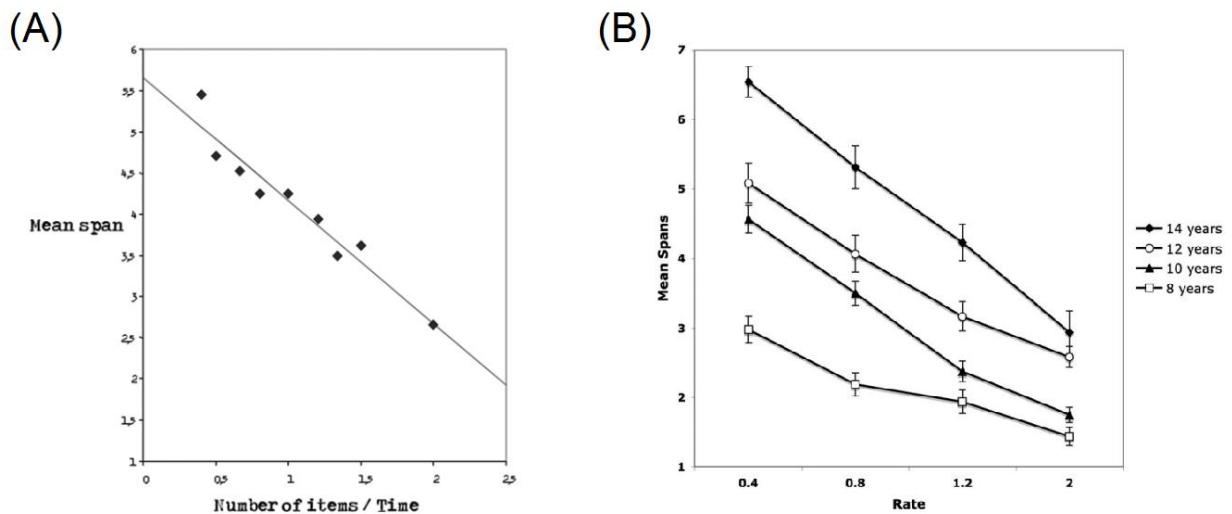


Figure 8. (A) Nombre moyen de stimuli retenus en mémoire de travail, en fonction du ratio entre le nombre d'items à traiter dans le cadre d'une tâche interférente et le temps disponible pour ce traitement (une valeur de 1 indique 1 item à traiter en 1 seconde). Tiré de Barrouillet, P., & Camos, V. (2010). Working memory and executive control: A time-based resource sharing account. *Psychologica Belgica*, 50(3&4), 353-382.

(B) Mêmes variables, représentées pour quatre groupes d'âges différents. Tiré de Barrouillet, P., et al. (2009). Working memory span development: A time-based resource-sharing model account. *Developmental Psychology*, 45(2), 477-490.

Une dernière thématique que nous avons explorée en lien avec le rôle du contrôle cognitif dans la mémoire de travail concerne le domaine du haut potentiel intellectuel. Les personnes à haut potentiel tendent à montrer une capacité en mémoire de travail nettement plus élevées que des participants contrôles (Rodríguez-Naveiras et al., 2019). Ce constat m'intéresse puisqu'il amène à se questionner sur l'origine de la corrélation entre mémoire de travail et performance complexe (Gonthier, 2016). Si l'on aborde la mémoire de travail sous

l'angle constructiviste, on peut faire l'hypothèse que le haut potentiel intellectuel va de pair avec un fonctionnement plus efficace de l'un des mécanismes qui contribuent à la capacité en mémoire de travail - mais lequel ? Cette question est d'intérêt à la fois pour la compréhension de la mémoire de travail et pour la compréhension des mécanismes en jeu dans le haut potentiel intellectuel. Le contrôle cognitif constitue un bon candidat, dans la mesure où les personnes à haut potentiel tendent également à faire preuve d'un contrôle cognitif plus efficace (e.g. Duan et al., 2009; Liu et al., 2011).

Dans un projet porté par l'équipe d'Alexandre Aubry et Béatrice Bourdin, nous nous sommes donc intéressés aux mécanismes pouvant être à l'origine de la supériorité en mémoire de travail des enfants à haut potentiel (Aubry, Gonthier, & Bourdin, en révision). À partir d'un échantillon d'enfants contrôles et à haut potentiel intellectuel ($QI > 125$), nous avons vérifié si la performance de contrôle cognitif (évalué à l'aide de l'ANT ; Fan et al., 2002) médiatisait le lien entre intelligence et capacité en mémoire de travail (évaluée à l'aide de l'ACCES). Nous avons observé une performance supérieure des enfants à haut potentiel aussi bien en mémoire de travail qu'en contrôle cognitif, mais aucune médiation du lien entre intelligence et mémoire de travail par le contrôle cognitif, contrairement à notre hypothèse.

En revanche, nous avons observé une médiation inattendue par la performance (vitesse et temps de réponse) sur les tâches concurrentes de traitement de l'information de l'ACCES : autrement dit, la plus forte capacité en mémoire de travail des enfants à haut potentiel s'explique en partie par un traitement plus rapide et plus efficace des tâches interférentes (juger si une phrase ou une opération mathématique est correcte, si une image est symétrique). Ce mécanisme va à l'encontre de notre hypothèse initiale portant sur un rôle du contrôle cognitif, mais rejoint une hypothèse ancienne sur le rôle des compétences tâches-spécifiques dans une tâche de mémoire de travail (e.g. Daneman & Carpenter, 1980). Cette hypothèse fait sens dans la perspective du modèle TBRS (un traitement plus efficace des tâches concurrentes laissant plus de temps libre pour le rafraîchissement de l'information), et constitue un mécanisme plausible (il est logique que les enfants à haut potentiel soient plus efficaces pour réaliser ce type de traitement de l'information). Des résultats préliminaires suggèrent également une médiation partielle par l'usage de stratégies, les enfants à haut potentiel ayant tendance à utiliser des stratégies de mémorisation un peu plus efficaces. Ce dernier point illustre l'intérêt des stratégies, qui font l'objet de la section suivante.

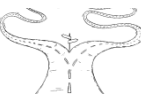


3.3. Le rôle des stratégies

Les stratégies peuvent jouer un rôle significatif dans les tâches de mémoire de travail. Dans le cas des tâches verbales, l'autorépétition (subvocale) des stimuli à mémoriser constitue un exemple classique et très étudié de comportement stratégique qui contribue directement à la performance. L'autorépétition n'est pas toujours considérée explicitement comme une stratégie et est souvent présentée comme un mécanisme structurel, faisant partie intégrante de la mémoire de travail verbale (e.g. Baddeley, 1986). Elle présente pourtant toutes les apparences d'un comportement stratégique : il s'agit d'une procédure bien identifiée, consciente ou conscientisable, que les participants peuvent choisir d'utiliser ou non, et qui influence qualitativement la façon dont la tâche est réalisée tout en affectant la performance. Il existe d'autres stratégies possibles en mémoire de travail verbale, par exemple le chunking des éléments à mémoriser ; et dans le cas de tâches portant sur des mots : regroupement en phrases, imagerie mentale (e.g. Dunlosky & Kane, 2007).

Ce tableau est assez clair en ce qui concerne la mémoire de travail verbal. Dans le cadre d'un projet utilisant une tâche complexe de mémoire visuo-spatiale (Thomassin, Gonthier, Guerraz, & Roulin, 2015), je me suis heurté à une question en apparence simple : quelles sont en fait les stratégies utilisées en mémoire à court terme visuo-spatiale ? Dans ce projet, nous avons initialement pensé évaluer directement les stratégies à l'aide d'un questionnaire mais nous n'avons pas réussi à déterminer précisément quelles stratégies étaient possibles dans une tâche de ce type. Cela m'a amené à constater le très peu de connaissance que nous avons, en général, sur la nature des stratégies possibles dans des paradigmes pourtant extrêmement classiques en mémoire à court terme ou en mémoire de travail visuo-spatiale - par exemple mémoriser des positions spatiales dans une grille. Cette ignorance est surprenante au regard des connaissances assez développées dont nous disposons quant aux stratégies dans les tâches de mémoire verbale : n'importe quel psychologue est capable d'évoquer quelques stratégies verbales comme l'autorépétition ou le chunking, mais lorsque l'on pose une question similaire au sujet d'une tâche visuo-spatiale, on se heurte souvent à un silence perplexe.

Ce questionnement m'a amené à réaliser une synthèse exhaustive de la littérature concernant les stratégies en mémoire à court terme / mémoire de travail visuo-spatiale (Gonthier, 2020), un effort rendu difficile par le fait que très peu d'études se sont intéressées explicitement à ce thème et que la quasi-totalité des stratégies identifiées apparaissent au



hasard d'expériences portant sur d'autres sujets. Les principales stratégies relevées sont résumées dans le Tableau 2. Ce travail de synthèse m'a amené à faire plusieurs constats. D'une part, les stratégies possibles dans une tâche visuo-spatiale sont nettement plus diversifiées qu'on ne pourrait s'y attendre - on peut relever plus d'une trentaine de stratégies courantes. Ces stratégies peuvent être regroupées (du moins conceptuellement) en sept processus stratégiques majeurs, qui incluent par exemple le chunking et l'autorépétition, mais aussi le recodage en mémoire ou l'encodage des positions spatiales relativement à d'autres éléments.

Tableau 2. *Processus stratégiques représentés dans les tâches de mémoire à court terme visuo-spatiales et leurs principales déclinaisons.*

Strategic process	Examples of implementations
1. Chunking	Memorize spatially close stimuli together Memorize temporally close stimuli together Memorize similar stimuli together Memorize geometric shapes Memorize meaningful shapes (letters...) Memorize symmetrical targets together
2. Holistic encoding	Encode all stimuli as a unitary picture Encode the shape of the path between stimuli Construct a story to describe the whole picture Keep the gaze fixed on the center of the screen
3. Relational encoding	Encode the position of stimuli in relation with the frame Encode the position of stimuli in relation with each other Encode the position of stimuli in relation with a landmark
4. Subdivision and categorical encoding	Divide the reference frame into sub-parts Divide a set of stimuli into sub-parts Divide a sequence of stimuli into sub-parts
5. Recoding	Verbal recoding : attach names to stimuli Verbal recoding : describe the display with a sentence Verbal recoding : use of system of coordinates Kinesthetic recoding : retrace a sequence of stimuli Kinesthetic recoding : retrace outlines of the shape Other recoding, such as auditory recoding
6. Long-term memory, semantic or episodic	Construct a meaningful story (holistic) Memorize meaningful configurations (chunking) Assign meaningful names to stimuli (verbal recoding) Use memories of configurations from previous trials
7. Visuo-spatial rehearsal	Rehearse the path between stimuli in a sequence Mentally refresh the visual image of the display
8. Others	Count the total number of stimuli Bind the stimuli in a sequence with digits Mental reduction of the encoding space Selective encoding of certain stimuli

Il ne semble pas exister de grandes différences d'efficacité entre ces processus stratégiques, qui paraissent tous pouvoir être bénéfiques à la performance (quoique les

données manquent pour conclure avec certitude sur ce point). La plupart des stratégies semblent toutefois aller de pair avec des patterns d'erreurs particuliers et des distorsions lors du rappel. Ainsi, la mémorisation holistique d'un ensemble de cibles comme un tout signifiant tend à générer un rappel plus proche d'une forme "idéale" que la disposition réelle des cibles ; l'encodage relationnel par rapport à l'un des bords de la grille d'encodage génère des erreurs plus importantes sur l'axe orthogonal à ce bord ; le chunking d'un ensemble de cibles peut générer un rappel totalement incorrect lorsque l'une des cibles est mal positionnée au rappel.



D'autre part, la littérature suggère qu'il n'existe pas de différences inter-individuelles marquées dans l'usage de ces stratégies, qui semblent plutôt varier de façon intra-individuelle en fonction des spécificités d'un essai : une stratégie d'élaboration sémantique par exemple doit s'appuyer sur une configuration spatiale particulière pour fonctionner, et ne peut donc être mise en place que lorsque la tâche et la disposition à mémoriser le permettent. Il est d'ailleurs particulièrement intéressant de constater que les stratégies listées ici ne sont pas exclusives : lors d'un même essai, un participant pourrait encoder un ensemble de positions spatiales sous la forme d'un chunk, situer ce chunk par rapport à un point de repère, lui donner une valeur sémantique en remarquant qu'il ressemble par exemple à une forme géométrique, puis auto-répéter les cibles à mémoriser par un mécanisme de rafraîchissement attentionnel.



Enfin, on remarque aussi que l'ensemble de ces processus stratégiques ont été relevés aussi bien pour des tâches de mémoire visuo-spatiales à présentation simultanée (l'ensemble des positions spatiales formant un pattern à mémoriser) qu'à présentation séquentielle (les cibles apparaissant l'une après l'autre) ; autrement dit, aussi bien pour des tâches visuelles que pour des tâches spatiales. Ce constat aboutit à questionner la distinction classique entre mémoire visuelle et mémoire spatiale, présentées comme des systèmes séparés (e.g. Logie, 2018) : il semble en fait exister une importante similarité fonctionnelle dans les mécanismes mis en jeu. On peut achever cette déconstruction de l'approche structurelle en remarquant que sur les sept principaux processus stratégiques identifiés, six se retrouvent à l'identique dans les tâches de mémoire de travail verbales et dans les tâches visuo-spatiales. Ce point parle à mon avis fortement en faveur de l'idée constructiviste que ces deux types de tâches ne correspondent pas à deux systèmes indépendants, mais plutôt à deux ensembles de processus fonctionnels qui se chevauchent largement, avec quelques processus spécifiques à l'un ou l'autre type de matériel (Majerus et al., 2010 ; Morey, 2018).



4. Mémoire de travail et cognition de haut niveau

J'ai évoqué le lien important entre capacité en mémoire de travail et cognition de haut niveau, notamment dans le cas de l'intelligence. Nous savons que la capacité en mémoire de travail est le meilleur prédicteur du facteur général de l'intelligence (Ackerman et al., 2005 ; Oberauer et al., 2005), à tel point que certains auteurs ont fusionné ces deux notions (Blair, 2006 ; Kyllonen & Christal, 1990). L'explication traditionnelle est une explication endogène : la mémoire de travail est un composant du système qui participe à la résolution des tests de haut niveau, et une capacité plus importante en mémoire de travail facilite d'une façon ou d'une autre l'intégration en mémoire des différents éléments de représentation qui doivent être combinés entre eux pour parvenir à une solution. L'analogie traditionnellement utilisée pour présenter cette idée est celle d'un tableau noir : la mémoire de travail est le tableau sur lequel on réalise les opérations cognitives, et un plus grand tableau autorise des opérations plus complexes.

Nous avons mené quelques expériences qui s'inscrivaient dans le cadre de ce type d'approche endogène, notamment dans le domaine de la prise de décision (Bagneux, Thomassin, Gonthier, & Roulin, 2013). Très peu de travaux ont examiné sur le lien entre mémoire de travail et prise de décision, comparativement à la plupart des autres activités cognitives de haut niveau. Nous nous sommes intéressés à cette question dans le cas particulier de l'*Iowa Gambling Task*, une tâche de prise de décision qui demande aux participants de tirer des cartes au sein de quatre piles (qui peuvent représenter des entreprises sur lesquelles investir ou bien un jeu de casino). Deux des piles sont avantageuses et deux sont désavantageuses, mais cette répartition n'est pas immédiatement apparente pour le participant : par exemple, la seconde pile génère un gain pour les neuf premières cartes piochées, mais une perte si importante pour la dixième qu'elle génère un déficit net.

L'*Iowa Gambling Task* est généralement décrite comme trop complexe pour permettre une analyse consciente ; la réussite est interprétée comme découlant de l'émergence progressive d'indices émotionnels inconscients (*hunches*) qui guident implicitement le participant dans sa prise de décision, ce qui laisse peu de place à un rôle de mécanismes cognitifs comme la mémoire de travail (Bechara et al., 1997, 1998). Mais la littérature donne en réalité des résultats contradictoires à ce niveau, les lésions du cortex dorsolatéral préfrontal donnant lieu ou non à des déficits et l'imposition d'une charge en mémoire de travail faisant ou

non chuter la performance (e.g. Hinson et al., 2002 ; Manes et al., 2002 ; Turnbull et al., 2005). À l'époque où nous avons réalisé ce travail, aucune étude n'avait testé l'effet des différences inter-individuelles en mémoire de travail dans un échantillon classique d'adultes sains, ce que nous avons fait.

Nos résultats (Bagneux et al., 2013) ont effectivement montré une corrélation entre capacité en mémoire de travail et réussite à l'*Iowa Gambling Task*, soutenant ainsi l'hypothèse d'un lien entre mémoire de travail et prise de décision. Cette tâche est plus particulièrement intéressante dans la mesure où elle renseigne sur l'évolution intra-individuelle de la performance : on sait en effet qu'il faut environ une quarantaine d'essais pour que les participants détectent quelles piles sont avantageuses et se mettent à prendre des décisions favorables. Nous avons observé que la capacité en mémoire de travail n'avait pas de lien avec la performance au cours des deux premiers blocs de 20 essais, et que la différence entre participants à faible et forte capacité en mémoire de travail n'émergeait qu'à partir du troisième bloc de 20 essais. Cet examen un peu plus fin qu'un score total permet une conclusion légèrement plus spécifique en faveur d'un rôle de la mémoire de travail en tant que telle (plutôt qu'une corrélation qui s'expliquerait, par exemple, par un engagement plus important dans la tâche de façon générale). Malheureusement, nous n'avons pas répliqué ces résultats dans une seconde étude, que nous ne sommes jamais parvenus à publier.

L'explication endogène du rôle de la mémoire de travail a l'avantage d'être intuitive mais pose quelques problèmes en pratique. D'une part, il est parfois difficile de définir exactement à quoi correspondent exactement les "représentations" qui doivent être intégrées entre elles pour résoudre une tâche de haut niveau, et dont le nombre justifie l'avantage dont disposent les participants à forte capacité en mémoire de travail. On peut l'imaginer pour les matrices de Raven ou une tâche de syllogismes, plus difficilement pour une tâche d'intelligence cristallisée. Dans le cas de l'*Iowa Gambling Task*, ces représentations à intégrer en mémoire pourraient être constituées par le pattern de gains et de pertes au cours des essais précédents, mais il est difficile de faire un lien direct entre la capacité en mémoire de travail d'un adulte typique (environ 4 items ; Cowan, 2001) et les patterns à intégrer en mémoire dans cette tâche au moment où la prise de décision avantageuse commence à apparaître (faut-il intégrer les caractéristiques de 4 piles de cartes ? des 40 essais précédents ?).



D'autre part, les études qui se sont intéressées à mettre réellement en lien le nombre de représentations requises par un problème donné et la capacité en mémoire de travail ont trouvé peu de liens entre les deux : par exemple, la capacité en mémoire de travail n'est pas nécessairement plus prédictive de la performance sur les syllogismes qui demandent l'intégration d'un plus grand nombre de modèles mentaux (Copeland & Radvansky, 2004), et la corrélation entre capacité en mémoire de travail et performance aux matrices de Raven ne dépend pas du nombre de règles logiques présentes dans une matrice (Salthouse, 1993 ; Unsworth & Engle, 2005).

Plus généralement, adopter une posture constructiviste rend forcément suspecte la corrélation entre deux construits comme la mémoire de travail et l'intelligence (Gonthier, 2016, 2017a ; Gonthier & Roulin, 2016) : si l'on voit la performance sur chacun de ces deux types de tâches comme résultant de l'opération d'un système constitué d'un grand nombre de mécanismes en interaction, alors la corrélation observée peut être due à un chevauchement dans n'importe lequel des mécanismes en jeu. Cette idée est depuis longtemps développée par les chercheurs s'intéressant au caractère prédictif de la mémoire de travail : les études menées dans la perspective du modèle d'Engle et collaborateurs (e.g. Engle et al., 1999) se sont largement appuyées sur l'hypothèse que la mémoire de travail correspond à la mémoire à court terme et au contrôle cognitif, et que la corrélation observée avec la cognition de haut niveau est due à un rôle commun du contrôle cognitif dans les deux tâches. Cette idée est aussi à la base de la *process overlap theory* (Kovacs & Conway, 2016).



Cette hypothèse intéressante se heurte à l'importante difficulté de mesurer le contrôle cognitif, la plupart des études argumentant en ce sens procédant par soustraction (e.g. Engle et al., 1999 ; Conway et al., 2002 ; Kane et al., 2004) : tester le lien entre mémoire de travail et intelligence en contrôlant la mémoire à court terme, avec l'idée que la corrélation partielle qui en résulte représente la contribution unique du contrôle attentionnel. Cette logique soustractive méconnaît bien sûr à la fois le rôle d'autres déterminants communs que le contrôle attentionnel (la motivation est encore une fois un exemple plausible), et le rôle du contrôle attentionnel dans la mémoire à court terme ; les résultats obtenus apparaissent de plus relativement instables (Colom et al., 2006 ; Unsworth & Engle, 2007b). Mais les études suivantes, même si elles ont souligné la contribution d'autres mécanismes communs au lien entre mémoire de travail et cognition de haut niveau, tels que la mémoire secondaire



(Unsworth & Engle, 2007a), ont conservé la logique constructiviste d'une approche par chevauchement de mécanismes communs comme explication à la corrélation observée.

Dans une autre série de travaux, nous avons exploré une hypothèse légèrement différente : le lien entre capacité en mémoire de travail et cognition de haut niveau pourrait être dû à l'usage de stratégies plus efficaces par les participants dotés d'une plus grande capacité en mémoire de travail. Comme dans le chapitre précédent, cette hypothèse présente le triple intérêt de proposer un mécanisme bien défini qui permet d'expliquer la plus grande réussite à un niveau de description individuel, plutôt que sous la forme d'une simple corrélation entre des variables définies au niveau inter-individuel ; de donner lieu à des prédictions plus spécifiques qu'un avantage générique pour les participants à forte capacité en mémoire de travail ; et si elle est vérifiée, de remettre en cause une position réaliste trop simpliste en montrant que des participants de niveaux d'aptitude différents ont des approches qualitativement différentes de la tâche, sur laquelle un continuum d'aptitude ne suffit donc plus à les classer.



Plusieurs études ont montré que les participants dotés d'une forte capacité en mémoire de travail tendent à utiliser des stratégies plus complexes dans les tâches de haut niveau (e.g. Beilock & Carr, 2005 ; Gimmig et al., 2006 ; Schelble et al., 2012). On remarque d'ailleurs que cette tendance subsiste même lorsqu'une stratégie complexe est moins efficace qu'une stratégie plus simple (Beilock & DeCaro, 2007), ce qui apporte à mon avis un argument fort en faveur d'un véritable lien entre mémoire de travail et tendance à s'engager dans une stratégie complexe, dans la mesure où ce pattern va à l'encontre de la performance habituellement toujours supérieure pour les participants à forte capacité en mémoire de travail (il s'agit du même argument que dans le cas de l'AX-CPT et du contrôle proactif ; voir page 28).

Un débat récurrent consiste à déterminer si cet usage différentiel de stratégies constitue une source de variance non pertinente (*strategy-as-noise*), ou bien si l'usage de stratégies contribue à la puissance prédictive des tâches de mémoire de travail, les participants à forte capacité en mémoire de travail ayant plus tendance à utiliser des stratégies efficaces qui les rendent plus efficaces à travers différents types de tâches (*strategy-as-mediator*). Dans l'ensemble, la littérature a plutôt soutenu l'hypothèse selon laquelle les stratégies sont une source d'erreur dans la mesure de la capacité en mémoire de travail : l'usage de stratégies efficaces dans une tâche de mémoire de travail corrèle peu avec d'autres construits (Bailey et

al., 2008 ; Dunlosky & Kane, 2007), et la corrélation entre performance en mémoire de travail et autres construits peut même augmenter lorsque l'on demande aux participants d'utiliser une stratégie particulière (Turley-Ames & Whitfield, 2003), ce qui suggère plutôt que les stratégies perturbent la précision de la mesure de la capacité en mémoire de travail.

Une limite majeure de cette conclusion a trait au fait que ces études ne se sont intéressées qu'à l'usage de stratégies *dans la tâche servant à mesurer la mémoire de travail*. Elles ne testent donc pas l'hypothèse alternative, légèrement différente, selon laquelle une forte capacité en mémoire de travail prédit l'usage de stratégies efficaces dans les tâches complexes, qui explique la performance élevée (e.g. Cokely et al., 2006). Autrement dit, les stratégies en mémoire de travail pourraient constituer une source d'erreur dans l'estimation de la capacité en mémoire de travail, tandis qu'en parallèle, les stratégies sur les tâches cognitives complexes sont déterminées par la mémoire de travail et contribuent à expliquer son rôle prédictif. Après tout, il s'agit de stratégies différentes, les stratégies en mémoire de travail n'étant pas transférables - par exemple - à une tâche d'intelligence, et les stratégies complexes évoquées dans la littérature pour certaines tâches pourraient bénéficier d'une capacité élevée en mémoire de travail.

Une solution possible pour tester cette hypothèse est de mesurer les stratégies utilisées sur la tâche complexe, et de vérifier si cette mesure des stratégies médiatise le lien entre capacité en mémoire de travail et performance ; mais cette approche présente l'habituelle limite interprétative de l'analyse de corrélations, et demande surtout une connaissance fine des stratégies possibles sur la tâche complexe étudiée. L'autre solution possible est de manipuler expérimentalement les stratégies ou la faisabilité des stratégies sur la tâche. C'est cette approche que nous avons retenue dans une étude qui avait pour objet de tester si les stratégies contribuent à la valeur prédictive de la mémoire de travail pour d'autres tâches mnésiques (Thomassin, Gonthier, Guerraz, & Roulin, 2015; Thomassin, Gonthier, & Roulin, 2012).

Notre démarche a consisté à combiner l'approche expérimentale et l'approche corrélationnelle (Cronbach, 1957). Le principe de cette démarche est d'évaluer la corrélation entre deux mesures, dans différentes conditions expérimentales créées par la manipulation directe d'une variable. Si cette manipulation expérimentale affecte la magnitude de la corrélation entre les deux mesures, on peut conclure que la manipulation expérimentale affecte un déterminant causal de la corrélation entre les deux mesures. Autrement dit, il s'agit



d'introduire une manipulation expérimentale dans l'analyse d'une corrélation, de façon à tester la causalité à l'origine de cette corrélation.

Dans le cas présent, nous avons relevé dans la littérature deux manipulations expérimentales interprétées comme interférant avec l'utilisation de stratégies efficaces de la part des participants à forte capacité en mémoire de travail : la double tâche (Schelble et al., 2012), et la contrainte temporelle imposée par la réduction des temps de présentation des stimuli (e.g. Friedman & Miyake, 2004 ; St Clair-Thompson, 2007). Après une évaluation de la capacité en mémoire de travail à l'aide d'un empan d'opérations, nous avons demandé à un ensemble de participants de réaliser une tâche complexe de mémoire à court terme (mémoriser des nuances de couleurs apparaissant dans une grille), en manipulant la présence ou non d'une double tâche et la présence ou non d'une contrainte temporelle forte. Notre hypothèse était que la capacité en mémoire de travail prédirait la performance sur la tâche complexe, mais uniquement dans la situation de simple tâche sans contraintes temporelles qui autorise la mise en place de stratégies efficaces.

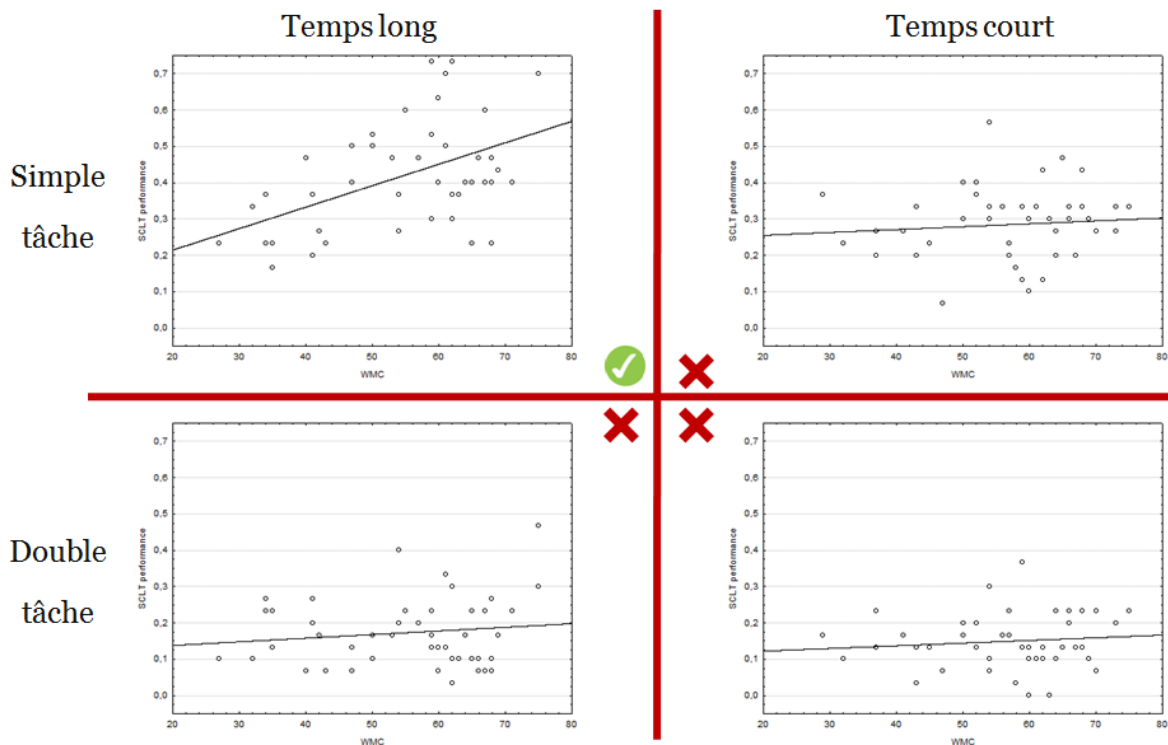


Figure 9. Corrélation entre capacité en mémoire de travail et performance sur une tâche complexe de mémorisation, en fonction de la condition expérimentale. Adapté de Thomassin, N., Gonthier, C., Guerraz, M., & Roulin, J.-L. (2015). The Hard Fall Effect: High working memory capacity leads to a higher, but less robust short-term memory performance. *Experimental Psychology*, 62(2), 89-97.

Les résultats sont présentés dans la Figure 9 et ont permis de valider cette hypothèse : la capacité en mémoire de travail ne prédit la performance que lorsque des stratégies efficaces sont possibles, et l'effet de la capacité en mémoire de travail interagit significativement avec la condition expérimentale. On remarque au passage que les difficultés supplémentaires imposées par la double tâche et la réduction du temps de présentation affectent très peu les participants à faible capacité en mémoire de travail, tandis qu'elles génèrent une chute importante pour les participants à forte capacité en mémoire de travail. Ce phénomène, déjà identifié par la littérature (e.g. Beilock & DeCaro, 2007), est aussi très proche de l'effet représenté dans la Figure 8 dans le contexte d'une approche développementale du modèle TBRS ; nous l'avons baptisé *Hard Fall Effect*.

Nos résultats montrent donc que la capacité en mémoire de travail n'est prédictive de la performance dans une tâche complexe que lorsque cette tâche permet la mise en place de stratégies efficaces. Nous avons obtenu des résultats similaires, non publiés, avec un autre type de tâche de mémorisation complexe (Thomassin, Gonthier, Guerraz, & Roulin, 2013). Dans l'ensemble, ces résultats sont compatibles avec l'hypothèse selon laquelle la capacité en mémoire de travail prédit la performance dans les tâches de haut niveau parce qu'une forte capacité en mémoire de travail va de pair avec l'utilisation de stratégies efficaces, et soulignent la relation qui existe entre capacité en mémoire de travail et comportements stratégiques, en particulier dans le domaine mnésique.

Bien sûr, cette conclusion n'est pas exempte de problèmes. La nature de la tâche complexe utilisée est discutable puisque proche d'une tâche de mémoire de travail, et l'on peut se demander si les résultats se généraliseraient, par exemple, à une tâche de raisonnement. En outre, les deux manipulations expérimentales retenues ici sont peu spécifiques : même si ce sont les plus généralement utilisées par la littérature, elles n'affectent pas *seulement* la mise en place de stratégies efficaces. Au vu de la Figure 8, on peut d'ailleurs se demander si l'interaction observée ne reflète pas plutôt un biais sous la forme d'un effet plancher pour les participants à faible capacité en mémoire de travail, comme je l'ai évoqué pour notre expérience portant sur le modèle TBRS. Mais cette expérience nous a permis de mettre à l'épreuve la démarche expérimentale-corrélacionnelle et a servi de base au premier projet présenté dans le chapitre suivant (page 128), dans le contexte de l'intelligence fluide.



4. Perspectives de recherche

4.1. Axe 1 : Outils d'évaluation de la capacité en mémoire de travail

Mon travail portant sur le développement de tâches d'évaluation de la mémoire de travail, mais aussi mon travail sur les stratégies en mémoire visuo-spatial, suggèrent quatre pistes de recherche possibles.

Perspective 2.1.A : le développement de l'ACCES a soulevé une question posée lors de l'expertise de l'article : une tâche adaptative est-elle "meilleure" qu'une tâche classique ? Il y a plusieurs façons d'entendre cette question - notamment "meilleure en terme de valeur prédictive pour d'autres construits", ou bien "meilleure en termes d'engagement du participant dans la tâche". Les arguments en faveur de l'acceptabilité d'une tâche adaptative pour le participant sont nombreux, mais ils sont de l'ordre du raisonnable et n'ont pas été testés empiriquement. Une comparaison entre deux tâches équivalentes, l'une adaptative et l'autre suivant une procédure ascendante classique, permettrait de répondre à cette question et pourquoi pas d'encourager l'adoption systématique de procédures adaptatives dans l'évaluation de la capacité en mémoire de travail. Une procédure élégante pour cela consisterait à entrelacer les versions adaptatives et non-adaptatives de différents subtests pour chaque participant, un même participant réalisant un subtest (par exemple verbal) adaptatif et un subtest (par exemple visuo-spatial) classique.

Perspective 2.1.B : les tâches du CCS et de l'ACCES tiennent compte de la variance spécifique au type de matériel, mais échouent à contrôler une autre source de variabilité dans les réponses : la variance méthode-spécifique, c'est-à-dire spécifique au paradigme d'empan complexe employé. Les empan complexes, en présentant en alternance une opération cognitive à réaliser et un stimulus à mémoriser, opérationnalisent bien la notion de mémoire de travail en associant explicitement mémorisation et manipulation de l'information ; ils constituent l'un des paradigmes les plus anciens (Case et al., 1979 ; Daneman & Carpenter, 1980) ; et sont de loin l'approche la plus utilisée pour évaluer les différences inter-individuelles en mémoire de travail (Conway et al., 2005 ; Redick et al., 2012). Mais les empan complexes sont loin d'être la seule épreuve possible pour évaluer la capacité en mémoire de travail.



Un autre exemple courant est la tâche de n -back (on présente une séquence de stimuli à l'écran, l'objectif étant de déterminer si le stimulus présenté est identique au stimulus présenté

n positions plus tôt), régulièrement utilisée en imagerie cérébrale. Le n -back est généralement considéré comme une "mesure de mémoire de travail", interchangeable avec un empan complexe ; pourtant la corrélation entre ces deux types de paradigmes est faible (Kane et al., 2007 ; Redick & Lindsey, 2013 ; mais voir aussi Schmiedek et al., 2014), ce qui suggère qu'il existe une part significative de variance liée à la méthode employée.

Il existe en fait au moins une dizaine de paradigmes différents pour évaluer la mémoire de travail :

1) Empans complexes pour lesquels l'information manipulée est différente de l'information à mémoriser. Exemples : empans de lecture, de symétrie et d'opérations (Unsworth et al., 2005) ; empans de rotation et de navigation (Kane et al., 2004) ; empans complexes avec activités élémentaires (Lucidi et al., 2014).

2) Empans complexes pour lesquels l'information à mémoriser est extraite de l'information à manipuler. Exemples : empan de lecture originel (Daneman & Carpenter, 1980) ; counting span (Case et al., 1979) ; tâche Mr. X (Alloway et al., 2010).

3) Tâches de "stockage dans le contexte d'un traitement" : contrairement aux empans complexes dans lesquels on alterne traitement et mémorisation pour chaque stimulus à mémoriser, il s'agit de présenter une seule liste de stimuli à mémoriser, puis une seule phase de traitement, avant le rappel. Exemples : Oberauer et al., 2003 ; Krumm et al., 2009.

4) Tâches de n -back : on présente une séquence de stimuli et le participant doit donner une réponse à chaque fois qu'un stimulus est identique au stimulus présenté n positions plus tôt, n pouvant aller de 1 à 3 ou 4, voire même 5 (e.g. Scharfen et al., 2018).

5) Mise à jour : on présente séquentiellement une liste de stimuli ou d'opérations ; le participant doit mettre à jour une ou plusieurs catégories à chaque nouvelle présentation. Exemples : keep-track task (Miyake et al., 2000 : on présente séquentiellement 15 mots appartenant à différentes catégories, et on demande au participant de rappeler le dernier mot présenté dans chaque catégorie) ; monitoring de sons (Miyake et al., 2000 : on présente des sons de différentes hauteurs, et on demande au participant d'indiquer lorsqu'il entend un son pour la quatrième fois) ; Paced Auditory Serial Addition Test ou PASAT (Naëgelé et Mazza, 2003 : on présente une séquence de chiffres, et le participant doit annoncer à chaque nouveau chiffre combien vaut la somme de ce chiffre et du précédent) ; running span (Broadway et

Engle, 2010 : on présente une séquence de stimuli de longueur imprévisible, et lorsque la séquence s'arrête le participant doit rappeler les k derniers stimuli présentés) ; Self-Ordered Pointing Task ou SOPT (Petrides et Milner, 1982 : on présente un ensemble d'images ; le participant doit pointer successivement chaque image, sans jamais pointer deux fois la même, et dans un ordre différent à chaque essai).

6) Traitement et mémorisation simultanés : on demande au participant d'effectuer des opérations successives sur un matériel donné ; les résultats intermédiaires doivent être mémorisés pour parvenir à la solution finale. Exemples : compteurs mentaux (Larson & Sacuzzo, 1989 : on incrémente ou décrémente séquentiellement différents compteurs et on demande au participant de rappeler la valeur finale de chaque compteur) ; mise à jour spatiale (Schmiedek et al., 2014 : on présente des points situés à différents emplacements, puis des déplacements à appliquer à ces points, et on demande au participant de rappeler l'emplacement final des points) ; soustractions en série du MMSE (en partant de 100, compter mentalement en soustrayant 7 ; Folstein et al., 1975), et plus généralement tâches d'arithmétique mentale ; réduction de symboles (Kyllonen et Christal, 1990 : recoder des séries de trois chiffres sous une forme condensée, par exemple en recodant deux chiffres consécutifs comme le même chiffre et deux chiffres différents comme le troisième chiffre : 1322131123 devient 22211 puis 231 puis 11 puis 1) ; raisonnement grammatical ABCD (Kyllonen et Christal, 1990 : "A précède B", "C ne précède pas D", "AB ne suit pas CD" : dans quel ordre sont les quatre lettres ?).

7) Coordination : la tâche consiste à coordonner plusieurs tâches ou opérations entre elles en gardant en mémoire l'état des différents éléments. Exemples : simulateur de contrôle aérien (Ackerman & Cianciolo, 2002) ; indiquer quand des points se déplaçant aléatoirement à l'écran forment des carrés ou quand des mots apparaissant aléatoirement à l'écran forment une rime (Oberauer et al., 2003).

8) Binding : la tâche consiste à mémoriser des éléments constitués d'une association entre plusieurs dimensions perceptives. Exemples : mémoriser des mots présentés dans une position spatiale particulière (Cowan et al., 2006), des couleurs présentées dans une position particulière (J. N. Wood, 2011), des formes particulières d'une couleur particulière à un emplacement particulier (Atkins et al., 2014).

9) Empans de transformation : la tâche consiste à mémoriser une séquence de stimuli, puis à transformer cette séquence avant rappel. Exemples : empan inverse (inverser la séquence

avant rappel) ; empan transposé (rappeler une séquence de chiffres après avoir appliqué une opération simple, telle que +2 ou -3, à tous les chiffres ; Fournet) ; recodage alphabétique (rappeler une série de lettres en décalant chaque lettre dans l'alphabet d'une position donnée, telle que +2 ; Kyllonen & Christal, 1990) ; empan alphabétique (rappeler une série de lettres ou de mots après les avoir remis dans l'ordre alphabétique ; Craik, 1986) ; séquence lettres-chiffres (rappeler une série de lettres et de chiffres après les avoir triés dans l'ordre croissant/alphabétique ; Wechsler, 2011) ; empan d'animaux (rappeler une série d'animaux dans l'ordre croissant de taille ; Schmiedek et al., 2014) ; tâche ABCD (rappeler une série d'éléments après les avoir remis dans un ordre spécifié à la fin par des instructions telles que "ail avant cèdre, chêne après sel" ; Kyllonen & Christal, 1990) ; empan conceptuel (rappeler dans le bon ordre les items d'une séquence, en rappelant uniquement les items d'une catégorie spécifiée à la fin de la séquence ; Kyllonen & Christal, 1990).

10) Reconnaissance en mémoire de travail : après présentation d'un ensemble de stimuli, le participant doit décider si une cible faisait partie de cet ensemble ou non. Exemples : tâche de S. Sternberg (1969) ; tâche de détection du changement (on présente un ensemble de cibles colorées, puis après un délai un ensemble identique ou légèrement différent, le participant devant détecter les éventuels changements ; Rouder et al., 2011).

L'importante hétérogénéité que l'on constate entre ces différentes tâches questionne directement l'idée que l'on mesure "la mémoire de travail", dans la mesure où elles impliquent des processus qui peuvent être assez différents. Cette idée est réminiscente de travaux dans des domaines très différents, comme par exemple l'étude de la dépression, où l'on observe un chevauchement assez faible entre les symptômes dépressifs listés par les différentes échelles d'évaluation de la dépression - ce qui aboutit à remettre en cause la réalité d'un concept unitaire de dépression (voir Fried, 2017). Dans le contexte de la mémoire de travail, ce constat ouvre une direction possible de recherche incluant la comparaison théorique des différentes tâches possibles en contrastant les processus théoriquement mis en jeu ; la comparaison empirique des tâches en évaluant leurs inter-corrélations et leur valeur prédictive pour d'autres construits ; ou encore la construction de batteries d'évaluation de la capacité en mémoire de travail qui font varier non seulement le type de stimuli (comme le CCS et l'ACCES), mais aussi le type de paradigme, de façon à construire une évaluation véritablement générale.

Perspective 2.1.C : mes travaux sur les stratégies et les patterns d'erreur qui leur sont associés m'ont conduit à m'interroger sur la façon dont on calcule les scores en mémoire de travail. La méthode la plus courante en psychologie différentielle consiste à calculer la proportion de stimuli rappelés dans la bonne position sérielle (Conway et al., 2005 ; Redick et al., 2012) : par exemple, pour la séquence 1-2-3-4-5, un participant qui rappelle 1-2-3-4 obtient 0.80 point. Mais on remarque que les erreurs d'intrusion, d'omission, de modification et d'inversion des stimuli n'ont pas toutes le même poids : ainsi, le rappel 1-2-4-5 n'obtient que 0.40 point puisque seuls les deux premiers éléments sont à la bonne position sérielle, tandis que le rappel 1-2-9-4-5 obtient 0.80 point. Par ailleurs, les erreurs d'intrusion et d'omission n'ont pas toujours le même poids en fonction de la position sérielle dans laquelle elles se présentent : le rappel 1-2-3-4 obtient 0.80 point, mais le rappel 2-3-4-5, quant à lui, obtient 0.00 point. Un problème identique se pose dans le cas des rappels de séquences de positions spatiales. Ce problème est d'autant plus ennuyeux que les erreurs d'omission, d'intrusion et d'inversion séquentielle n'ont pas toutes la même fréquence, et que leur fréquence dans une position sérielle donnée peut dépendre du niveau d'aptitude (Unsworth & Engle, 2006). Il existe certainement de meilleures méthodes de cotation : une méthode possible serait par exemple la *distance d'édition* (algorithme de Damerau-Levenshtein), qui correspond tout simplement au nombre de transformations requises pour que la séquence rappelée soit identique à la séquence-cible. Reste à examiner les propriétés psychométriques des solutions alternatives.



Par ailleurs, les travaux portant sur l'analyse des erreurs dans les tâches visuo-spatiales (et montrant par exemple un rappel biaisé en direction des angles de la grille de présentation lorsque les participants utilisent une stratégie consistant à diviser la grille en quatre quadrants) suggèrent une évolution possible pour ce type de tâches. La version archétypale est une grille dans laquelle on "allume" certaines cases pour constituer un pattern à mémoriser ou une séquence de cases à retenir. Mais rien n'impose en réalité un tel découpage de l'espace sous la forme d'une grille. On peut tout aussi bien présenter les localisations visuo-spatiales dans un espace continu, sans délimitation, comme l'ont fait les chercheurs intéressés par l'analyse des distorsions lors du rappel. Le participant devrait alors, pour la phase de rappel, déplacer librement son curseur à l'écran et cliquer sur les emplacements qui lui paraissent les plus proches des positions présentées à l'encodage. Cette approche donne lieu à une autre façon de concevoir la performance en mémoire visuo-spatiale : non pas comme la somme des éléments

rappelés correctement, mais comme la somme des distances angulaires entre les positions présentées à l'encodage et les positions restituées lors du rappel.

Autrement dit, il s'agit de penser la performance en mémoire de travail en termes de précision plutôt que de capacité. Le récent développement de modèles conceptualisant la mémoire de travail sous la forme d'une ressource allouée de façon continue à différents items, plutôt que d'une capacité discrète, correspond bien à cette idée (Ma et al., 2014). Ces deux approches ne sont d'ailleurs pas incompatibles, et on peut concevoir un modèle qui tient compte à la fois du caractère oublié ou non de la cible, et lorsque la cible n'est pas oubliée, de la précision du rappel (Rhodes et al., 2020).



Les premiers travaux formulés en termes de précision en sont à leurs balbutiements et il est difficile d'anticiper précisément la façon donc ce paradigme pourrait modifier l'évaluation des différences inter-individuelles. La représentation du construit serait légèrement différente ; des dissociations pourraient apparaître (Rhodes et collaborateurs suggèrent par exemple que le vieillissement engendre plus une baisse de précision du rappel qu'un oubli) ; ce format pourrait limiter l'usage de certaines stratégies (comme le recodage verbal, les systèmes de coordonnées...) et donc éliminer une source possible de variabilité inter-individuelle (dont il est incertain qu'elle soit en lien avec le niveau d'aptitude) ; ce format pourrait peut-être améliorer les propriétés psychométriques de la tâche, la mesure devenant continue et non plus discrète, et moins exposée aux difficultés de cotation telles que celles décrites ci-dessus.

Perspective 2.1.D : le clinicien francophone dispose de peu d'options pour évaluer la capacité en mémoire de travail. L'implication de déficits en mémoire de travail dans certains troubles développementaux, par exemple la dyslexie, la dyspraxie (e.g. Jeffries & Everatt, 2004), ou encore le TDA/H (Martinussen et al., 2005), mais aussi dans le cas de lésions cérébrales telles que les accidents vasculaires cérébraux (van der Linden et al., 2007) crée pourtant un intérêt à ce niveau chez les psychologues de l'enfant et les neuropsychologues. Pour la plupart des psychologues, l'évaluation semble passer généralement par l'indice de mémoire de travail des échelles de Wechsler. Il existe d'autres solutions plus marginales, comme l'empan d'opérations (Fournet et al., 2012) ou le PASAT (Naëgelé & Mazza, 2003).

La plupart de ces options sont payantes et/ou ne disposent pas d'un étalonnage. La solution d'utiliser l'IMT calculé à partir d'une échelle de Wechsler apparaît comme étant très peu satisfaisante en pratique, puisque cet indice s'appuie sur trois subtests dont deux sont en

réalité des tâches de mémoire à court terme et non des tâches de mémoire de travail. L'empan de chiffres endroit n'impose pas de demande significative de traitement de l'information et constitue donc bien sûr une tâche de mémoire à court terme ; quant à l'empan envers, plusieurs travaux ont montré que cette tâche sature un facteur de mémoire à court terme plus que de mémoire de travail (e.g. Conway et al., 2005 ; St Clair-Thompson, 2010), soit que le traitement de l'information demandé par l'inversion d'une séquence de stimuli constitue une demande trop faible pour interférer avec le rafraîchissement de la trace mnésique, soit que cette manipulation n'interfère pas avec le rafraîchissement parce qu'elle porte sur le matériel à mémoriser lui-même. Dans l'ensemble, il existe donc une demande pour une tâche évaluant la capacité en mémoire de travail qui s'appuie sur les données de la recherche.

Le développement d'une tâche de mémoire de travail à destination clinique soulève quelques réflexions. D'une part, la procédure adaptative semble particulièrement adaptée à cet usage, pour les raisons évoquées plus haut (page 81), dans la mesure où elle permet l'évaluation de patients qui présentent absolument n'importe quel niveau d'aptitude. D'autre part, les empan complexes traditionnels sont peu adaptés car très dépendants de l'aptitude à réaliser le traitement d'informations souvent complexes, comme le reflètent nos résultats dans le cadre du haut potentiel ; or, un patient en contexte clinique est justement susceptible de présenter un déficit précisément dans ces domaines d'aptitude (dyslexie, dyscalculie...), et plus généralement la performance à ce type de tâches interférentes est susceptible d'augmenter fortement avec l'âge sans que cela reflète forcément un gain dans la capacité en mémoire de travail en-dehors de ce type de paradigme particulier. L'utilisation de tâches interférentes très simples, comme le font les épreuves développées dans le cadre du modèle TBRS (Lucidi et al., 2014), semble nettement plus adaptée à cet usage. La tâche clinique développée devrait idéalement combiner plusieurs subtests verbaux et visuo-spatiaux, comme c'est le cas du CCS et de l'ACCES, pour minimiser la variance spécifique au type de matériel. Il restera à sélectionner le paradigme le plus approprié et le matériel le moins sensible aux troubles développementaux, puis à l'étalonner.



Perspective 2.1.E : le phénomène que nous avons baptisé hard fall effect, c'est-à-dire la chute de performance plus importante pour les participants à fort niveau d'aptitude quand on les place en situation de double tâche (Thomassin et al., 2015), rejoint des conclusions obtenues dans d'autres contextes (par exemple Barrouillet et al., 2009 ; Beilock & DeCaro, 2007 ; Schelble

et al., 2017) et illustre la généralité de ce pattern de résultats. Ce phénomène justifie un certain scepticisme au sujet d'un paradigme classique d'évaluation de la "mémoire de travail" : la double tâche utilisée par Baddeley et collaborateurs (Baddeley et al., 1996), et plus particulièrement le calcul d'un score d'efficacité de l'administrateur central comme la chute de performance entre simple tâche et double tâche.

Cet indicateur basé sur la chute de performance, appelé *mu*, peut se justifier logiquement, et semble fonctionner pour des adultes âgés atteints de démence qui démontrent une chute de performance catastrophique en double tâche (Baddeley et al., 1996). Mais il y a fort à parier qu'il atteint l'inverse de l'objectif escompté quand on l'utilise par exemple avec une population de jeunes adultes, ceux qui bénéficient de la meilleure capacité en mémoire de travail étant les plus susceptibles de chuter en double tâche (leur performance en baseline étant plus élevée et la double tâche interférant avec les procédures qu'ils utilisent pour étayer leur performance). L'usage continu de la double tâche et de l'indicateur *mu* en clinique neuropsychologique (voir par exemple Baddeley et al., 1997 ; Della Sala et al., 2010) pourrait donc être fortement remis en cause. Cette idée peut être testée simplement en calculant le score *mu* pour des participants dont on a estimé en parallèle le niveau d'aptitude en mémoire de travail ou en contrôle cognitif ; on peut s'attendre à une corrélation entre niveau d'aptitude au contrôle attentionnel (tel qu'estimé en-dehors du contexte de la double tâche) et chute de performance de la simple tâche à la double tâche.



4.2. Axe 2 : Déterminants stratégiques de la performance en mémoire de travail

Perspective 2.2.A : mon travail de synthèse sur les stratégies en mémoire de travail visuo-spatiale (Tableau 2) a le mérite d'offrir une première approche de la diversité des processus stratégiques possibles, mais ne répond que partiellement à la question de la variabilité des stratégies et de leur contribution à la performance. Ces thématiques ont été très peu explorées par la littérature : peu d'études ont mesuré les stratégies ; quelques-unes ont comparé la performance obtenue en utilisant une stratégie en particulier, sans s'intéresser aux autres. De plus, aucune étude ou presque n'a évalué les stratégies utilisées essai par essai, se contentant à la place d'un report global en fin de tâche. Le constat selon lequel les stratégies utilisées semblent dépendre de l'organisation des stimuli spécifique à un essai, plus que de différences inter-individuelles, parle pourtant fortement en faveur d'une analyse au niveau de l'essai plutôt qu'au niveau de la tâche ; en l'absence de stabilité inter-individuelle forte, ce niveau d'analyse est de toute façon le plus adapté dans une perspective constructiviste s'intéressant au mécanisme de la réussite dans une situation donnée. Une mesure des stratégies essai par essai permettrait aussi d'analyser plus finement l'existence de co-occurrences entre les stratégies.

J'ai mené un premier travail en ce sens, dont les résultats préliminaires ont fait l'objet d'une communication (Gonthier, 2018) : j'ai observé qu'un participant donné utilise en moyenne 12 stratégies différentes à travers la tâche, avec une moyenne de 3 stratégies par essai. Mais ces résultats doivent encore être analysés plus finement, notamment à l'aide d'analyses en modèles mixtes qui permettront de quantifier la part de variance dans l'utilisation de stratégies qui relève de l'individu et celle qui relève d'une configuration d'items donnée (Gonthier, en préparation). Peaufiner un protocole expérimental qui permette l'évaluation des processus stratégiques pourra ouvrir de nouvelles perspectives de recherche, notamment dans l'évaluation de la variabilité développementale dans les comportements stratégiques (Monnier & Gonthier, en préparation) : ce thème a été exploré dans le domaine de la mémoire de travail verbale, notamment autour du développement des stratégies d'autorépétition (Hitch et al., 1989), mais très peu dans le domaine de la mémoire de travail visuo-spatiale.

Perspective 2.2.B : lorsqu'il existe des stratégies efficaces et des stratégies peu efficaces, comme cela semble être le cas en mémoire de travail verbale (e.g. Dunlosky & Kane, 2007), une question se pose de façon récurrente : par définition, l'usage d'une stratégie efficace corrèle



avec une performance élevée, mais laquelle de ces deux variables est la cause de l'autre (Gonthier & Roulin, 2019) ? Autrement dit, l'usage d'une stratégie efficace est-il la cause ou la conséquence d'une performance élevée ? Les deux options sont plausibles : soit le fait d'utiliser une stratégie efficace crée une performance élevée sur la tâche, qui amène à la conclusion que le participant a une "aptitude élevée" ; soit le fait d'avoir une aptitude élevée permet au participant de sélectionner et d'utiliser la stratégie la plus efficace, et l'aptitude est donc la cause du comportement stratégique. On peut d'ailleurs imaginer une causalité circulaire : un participant d'aptitude élevée est celui qui est capable d'utiliser la stratégie la plus efficace, et cette stratégie efficace est ce qui détermine sa performance élevée.



Le rôle causal des stratégies et leur lien avec les différences inter-individuelles est d'intérêt pour tout chercheur qui s'intéresse à la variabilité des comportements stratégiques, notamment parce qu'il s'agit d'un élément qui peut aider à trancher le débat concernant l'intérêt des stratégies dans la mesure (la variabilité stratégique est-elle une source de variance d'erreur ou un contributeur important à la performance mesurée ?). Par ailleurs, ce rôle causal des stratégies peut être testé dans une perspective expérimentale-corrélacionnelle, simplement en manipulant les stratégies utilisées. Notre étude sur le hard fall effect (Thomassin et al., 2015) en est un exemple, mais on peut aussi manipuler plus directement le comportement stratégique en donnant explicitement des consignes en ce sens aux participants (e.g. McNamara & Scott, 2001 ; Turley-Ames & Whitfield, 2003).



Si une aptitude élevée cause de façon secondaire l'emploi de stratégies efficaces, alors demander aux participants d'utiliser une stratégie efficace ou inefficace ne devrait pas radicalement altérer le lien entre capacité en mémoire de travail et performance. Si les stratégies causent une performance élevée, alors induire l'usage d'une stratégie efficace devrait faire augmenter la performance des participants à faible capacité en mémoire de travail, tandis qu'induire l'usage d'une stratégie inefficace devrait faire chuter la performance des participants à forte capacité en mémoire de travail. Enfin, si aptitude et stratégie constituent une boucle causale ; autrement dit, si une aptitude élevée conditionne la possibilité d'utiliser une stratégie efficace, l'usage de cette stratégie efficace étant à l'origine de la performance plus élevée ; alors induire une stratégie efficace ne devrait pas augmenter la performance des participants à faible capacité en mémoire de travail, mais induire une stratégie inefficace devrait faire chuter la performance des participants à forte capacité en mémoire de travail.

5. Publications et communications associées

Articles publiés et en révision

Aubry, A., **Gonthier, C.**, & Bourdin, B. (en révision). Explaining the high working memory capacity of gifted children: Contributions of processing skills and executive control. *High Ability Studies*.

Bagneux, V., Thomassin, N., **Gonthier, C.**, & Roulin, J.-L. (2013). Working memory in the processing of the Iowa Gambling Task : An individual differences approach. *PLoS ONE*, 8(11) : e81498. doi : 10.1371/journal.pone.0081498

Gonthier, C. (2020). Charting the diversity of strategic processes in visuo-spatial short-term memory. *Perspectives on Psychological Science*.

Gonthier, C., Aubry, A., & Bourdin, B. (2017). Measuring working memory capacity in children using adaptive tasks: Example validation of an adaptive complex span. *Behavior Research Methods*. doi:10.3758/s13428-017-0916-4

Gonthier, C., & Gavornikova-Baligand, Z. (en révision). Are individual differences in working memory capacity grounded in the reactivation of memory traces ? *Learning and Individual Differences*.

Gonthier, C., & Hot, P. (2013). Apports de l'électro-encéphalographie à la compréhension de la mémoire. *Revue de neuropsychologie*, 5(4), 243-254. doi:10.1684/nrp.2013.0280

Gonthier, C., Thomassin, N., & Roulin, J.-L. (2016). The Composite Complex Span : French validation of a short working memory task. *Behavior Research Methods*, 48(1), 233-242. doi:10.3758/s13428-015-0566-3

Thomassin, N., **Gonthier, C.**, Guerraz, M., & Roulin, J.-L. (2015). The hard fall effect: high working memory capacity leads to a higher, but less robust short-term memory performance. *Experimental Psychology*, 62(2), 89-97. doi:10.1027/1618-3169/a000276

Articles en préparation

Gonthier, C. (en préparation). Strategy use in visuo-spatial short-term memory: dynamic selection in a pool of facilitatory processes.

Monnier, C., & **Gonthier, C.** (en préparation). Development of strategic processes in visuo-spatial short-term memory.

Communications

Aubry, A., **Gonthier, C.**, & Bourdin, B. (2017, Novembre – communication affichée). *Évaluer la mémoire de travail de l'enfance à l'âge adulte : Exemple d'une tâche adaptative*. 11èmes Journées Internationales de Psychologie Appliquée, Amiens, France.

Gonthier, C. (2016, Novembre - communication orale). *Vous avez dit mémoire de travail ?* Séminaire invité, 9ème journée du CRTLA de Rennes, Rennes, France.

Gonthier, C., (2016, Janvier – communication orale). *La place de la mémoire de travail dans la cognition : Multiples visages de la capacité en mémoire de travail*. Séminaire invité, CRP-CPO, Université de Picardie Jules-Verne, Amiens, France.

Gonthier, C., (2017a, Mars – communication orale). *Vers une approche multi-déterministe de la cognition de haut niveau*. Séminaire invité, LATI, Université Paris Descartes, Paris, France.

Gonthier, C., Aubry, A., & Bourdin, B. (2016, Juin – communication affichée). *Mesurer la mémoire de travail chez l'enfant : validation d'un empan complexe adaptatif*. XXIIèmes Journées Internationales de Psychologie Différentielle, Chambéry, France.

Gonthier, C., Baciú, M., & Roulin, J.-L. (2018 – actes de colloque). Approche différentielle et neuroimagerie : l'exemple des différences inter-individuelles en mémoire de travail. In J. Collange, E. Loarer, & T. Lubart (Eds.), *La psychologie différentielle au 21ème siècle : nouvelles modélisations et applications* (pp. 17-21). Paris, France: Editions Laboratoire LATI. ISBN:978-2-9563911-0-4

Gonthier, C., & Roulin, J.-L. (2016, Juin – communication orale). *La mémoire de travail n'existe pas ? Une approche multi-déterministe de la cognition de haut niveau*. XXIIèmes Journées Internationales de Psychologie Différentielle, Chambéry, France.

Gonthier, C., Thomassin, N., Baciú, M., & Roulin, J.L. (2013, Mai – communication affichée). *Est-on plus sensible au contexte de la tâche quand on a une bonne mémoire de travail ?* Journée scientifique de l'Ecole Doctorale Ingénierie pour la Santé, la Cognition et l'Environnement, Grenoble, France.

Gonthier, C., Thomassin, N., Baciú, M., & Roulin, J.-L. (2014, Juin – communication orale). *Une approche différentielle en neurosciences : différences inter-individuelles en mémoire de travail et imagerie cérébrale du contrôle cognitif*. XXIèmes Journées Internationales de Psychologie Différentielle, Paris, France.

Gonthier, C., Thomassin, N., & Roulin, J.-L. (2012, Juin – communication affichée). *Différences interindividuelles en mémoire de travail et mode de contrôle cognitif*. XXèmes Journées Internationales de Psychologie Différentielle, Rennes, France.

Thomassin, N., **Gonthier, C.**, Guerraz, M., & Roulin, J.L. (2013, Mai – communication affichée). *Quand les indices n'aident pas : l'effet pervers des indices de rappel est-il plus marqué*

avec une bonne stratégie de mémorisation ? Journée scientifique de l'Ecole Doctorale Ingénierie pour la Santé, la Cognition et l'Environnement, Grenoble, France.

Thomassin, N., **Gonthier, C.**, & Roulin, J. (2012, Juin – communication affichée). *Stratégies élaboratives et capacité de mémoire de travail*. XXèmes Journées Internationales de Psychologie Différentielle, Rennes, France.

Thèse de doctorat

Gonthier, C. (2014). *Cognitive control in working memory: An individual differences approach based on the Dual Mechanisms of Control framework* [Thèse de doctorat non publiée]. Département de Psychologie, Université de Grenoble.

CHAPITRE 3. INTELLIGENCE FLUIDE

Résumé

Une petite revue de la littérature sur les conceptions de l'intelligence est l'occasion de montrer que la position réaliste a été très peu représentée dans ce domaine, la grande majorité des chercheurs adoptant une orientation constructiviste quoique sans toujours en tirer toutes les conséquences. Des modèles fonctionnels de la performance existent depuis au moins une cinquantaine d'années.

Mes travaux dans le domaine de l'intelligence se sont intéressés à deux types de mécanismes contribuant à la performance : 1) Des mécanismes qualitatifs, sous la forme des stratégies de raisonnement, en lien notamment avec la capacité en mémoire de travail ; 2) Des déterminants plus quantitatifs, portant notamment sur le rôle du temps de réponse et des connaissances déclaratives. Ces travaux me conduisent à proposer en perspective un modèle intégratif de la multi-détermination des scores d'intelligence, avec en tête la possibilité d'une application clinique.

1. Revue de littérature

Les scores d'intelligence constituent un excellent prédicteur des mesures humaines dans un grand nombre de domaines, tels que les temps de réponse (Jensen, 1987) et la mémoire de travail (Ackerman et al., 2005) ; mais aussi la performance scolaire (Deary et al., 2007), le bonheur (Ali et al., 2013), le revenu et le prestige de la profession (Strenze, 2007), la taille (Pearce et al., 2005), le succès au travail (Kuncel et al., 2004), ou encore l'espérance de vie (Batty et al., 2009). Cette impressionnante puissance prédictive justifie l'intérêt des chercheurs et des cliniciens pour l'intelligence (quoique comme dans le cas de la mémoire de travail, ces

corrélations peu différenciées puissent être vues comme un problème de validité divergente qui questionne la nature de la ou des entités à l'origine du score d'intelligence).



Les enseignements qui portent sur l'intelligence en Licence de psychologie adoptent presque toujours une perspective historique, en commençant par les travaux fondateurs de Spearman (1904) : on présente aux étudiants le constat initial selon lequel l'analyse factorielle des épreuves d'intelligence fait émerger un facteur général, nommé *g*. L'existence de ce facteur *g* est ensuite utilisée comme argument en faveur de l'existence d'une aptitude générale de haut niveau, que l'on appelle intelligence, et qui est à l'origine de la performance sur les tests d'intelligence - quels qu'ils soient, et ce en dépit de leurs profondes différences de structure (hypothèse de l'indifférence de l'indicateur). Tel qu'on le présente aux jeunes psychologues, le facteur *g* semble ainsi refléter une aptitude unitaire à l'intelligence, qui joue directement un rôle causal dans la détermination de la performance (e.g. Gottfredson et al., 1997a: "*Differences in intelligence certainly are not the only factor affecting performance in education, training, and highly complex jobs (no one claims they are), but intelligence is often the most important*"). Cette idée est appuyée à la fois par le langage causal utilisé ("un score d'intelligence élevé entraîne..." ; "le facteur général explique la variance...") et par le type de modèles statistiques employés, qui présentent systématiquement le facteur général sous la forme d'une variable latente réflective.

Les recherches postérieures à Spearman ont certes mis en évidence l'existence de facteurs moins généraux correspondant à des regroupements entre sous-catégories de tests d'intelligence, selon la nature des aptitudes mises en jeu, verbales ou visuo-spatiales par exemple (Thurstone, 1935). Mais ces facteurs de niveau inférieur tendent à être eux-mêmes corrélés (Eysenck, 1939 ; voir Jensen, 1987), et les travaux ultérieurs utilisant l'analyse factorielle hiérarchique, qui ont abouti au relatif consensus actuel concernant le modèle CHC de l'intelligence (McGrew, 2009), conservent un facteur *g* au sommet de la hiérarchie factorielle des aptitudes. La vision du facteur *g* comme reflet d'une aptitude unitaire à l'intelligence qui serait à l'origine de la performance sur tous les tests ne semble donc pas sérieusement remise en cause, même si elle est tempérée par la reconnaissance d'une certaine hétérogénéité entre les domaines d'aptitude.

On voit bien à l'œuvre ici la réification implicite de l'intelligence, et le glissement épistémologique vers une position réaliste qu'entraîne cette réification : puisque les tests

d'intelligence corrèlent entre eux, ils ont une cause commune, cette cause s'appelle intelligence, et elle est représentée par le facteur g qui constitue une entité réaliste. Cette idée simple (une entité qui correspond à l'intelligence \rightarrow cause \rightarrow la performance sur les tests d'intelligence, et donc le facteur g identifié par Spearman) est si familière aux psychologues qu'il est amusant de constater à quel point elle constitue une simplification grossière de la position de Spearman lui-même à ce sujet. Même si Spearman est intéressé par la recherche de la cause du regroupement factoriel entre les tests d'intelligence, il prend grand soin de distinguer le facteur g (l'entité statistique) de l'intelligence (l'entité causale supposée être à son origine). Il écrit ainsi :

G means a particular quantity derived from statistical operations. Under certain conditions the score of a person at a mental test can be divided into two factors, one of which is always the same in all tests, whereas the other varies from one test to another; the former is called the general factor or G, while the other is called the specific factor. This then is what the G term means, a score-factor and nothing more. (Spearman, 1931 ; voir Deary et al., 2008)

Notice must be taken that this general factor g , like all measurements anywhere, is primarily not any concrete thing but only a value or magnitude. Further, that which this magnitude measures has not been defined by declaring what it is like, but only by pointing out where it can be found. It consists in just that constituent - whatever it may be - which is common to all the abilities inter-connected by the tetrad equation. This way of indicating what g means is just as definite as when one indicates a card by staking on the back of it without looking at its face. Such a defining of g by site rather than by nature is what was meant originally when its determination was said to be only "objective." Eventually, we may or may not find reason to conclude that g measures something that can appropriately be called "intelligence." Such a conclusion, however, would still never be a definition of g , but only a "statement about it." (Spearman, 1927)

Ces deux passages reflètent la nuance apportée par Spearman quant à l'interprétation des résultats apportés par la méthode de l'analyse factorielle, dont il est à l'origine : l'entité statistique qu'est g ne doit pas être confondue avec l'entité causale qui la provoque. Dans la mesure où le facteur g représente simplement la corrélation générale observée entre les épreuves d'intelligence, cette distinction entre le facteur statistique et les déterminants qui en

sont à l'origine n'est ni plus ni moins qu'une paraphrase de l'antienne selon laquelle une corrélation n'est pas une causalité.

La prudence de Spearman quant à la nature de l'entité causale à l'origine de la corrélation observée sous la forme du facteur *g* se retrouve en fait chez une grande partie des théoriciens de l'intelligence au XX^{ème} siècle. En réponse à un psychologue qui critique la réification du facteur *g* en une entité sous-jacente, Arthur Jensen, pourtant l'un des plus ardents défenseurs de l'étude de *g* en tant que tel, répond : *No one who has worked in this field has ever thought of g as an entity or thing. Spearman, who discovered g, actually said the very same thing that you're saying now, and Cyril Burt and Hans Eysenck said that also: just about everyone who has worked in this field has not been confused on that point.* (voir Bock et al., 2000). Il existe de fait d'excellentes discussions du statut ontologique des facteurs issus d'analyses factorielles dans le contexte de l'intelligence, et ce depuis longtemps (e.g. Burt, 1940, *The logical and metaphysical status of factors in psychology*).

Il est donc clair pour les théoriciens de l'intelligence, à défaut des psychologues en général, que le facteur *g* ne reflète pas l'opération d'une entité concrète. Alors que reflète-t-il ? Par définition, le facteur *g* reflète ce qui provoque la corrélation entre différents tests ; autrement dit, *ce qui est à l'origine de la part de variabilité inter-individuelle de performance qui est commune à un ensemble de tests d'intelligence, dans un échantillon donné.* On peut admettre que cette variation commune à plusieurs tests est bel et bien provoquée par *quelque chose* ; la question centrale soulevée par cette définition est alors celle du nombre de déterminants qui sont à l'origine des corrélations observées. Rien n'indique que la corrélation entre deux tests est due à un unique déterminant qui agit sur ces deux tests ; il est tout aussi valable de supposer qu'il existe deux déterminants qui en moyenne agissent à l'identique sur ces deux tests, ou trois déterminants, ou plus. Si l'on imagine que la performance à n'importe quel test d'intelligence est simplement le produit de l'intelligence et de la motivation, alors ces deux facteurs influencent la performance sur tous les tests, ils contribuent conjointement à organiser la variabilité inter-individuelle, et la corrélation entre les tests est causée simultanément par ces deux fonctions mentales. On peut donc très bien imaginer un facteur général d'origine multi-déterminée.



On doit la première démonstration de ce point à Thorndike et à Thomson, à l'origine de la *sampling theory* du facteur général (voir Jensen, 1998). Cette théorie propose que la

performance aux tests d'intelligence pourrait être due à un très grand nombre de déterminants, dont certains sont communs à plusieurs tests et d'autres spécifiques à un test ; une corrélation entre deux tests serait due simplement à un chevauchement partiel entre les déterminants impliqués par ces deux tests. Pour illustrer cette idée, Thomson (1916) génère aléatoirement des scores pour un ensemble de tests d'intelligence fictifs, en additionnant 145 déterminants de la performance (des dés à 6 faces), dont 36 sont communs à tous les tests. Thomson obtient avec ces données un facteur général et reproduit les résultats de Spearman. Des travaux plus récents ont confirmé que les modèles de Spearman et de Thomson s'ajustent tout aussi bien aux mêmes données d'un point de vue statistique (Bartholomew et al., 2009, 2013) et donnent donc lieu à des explications du facteur général qui ne peuvent se distinguer sur la base d'analyses factorielles.

C'est à ce point du raisonnement que pourraient commencer les querelles théologiques au sujet du caractère uni- ou multi-déterminé du facteur g , qui seraient nécessairement de l'ordre de la foi puisque les données empiriques ne permettent pas de trancher. Étonnamment, il est en fait difficile de présenter l'état de la controverse, car il n'y a pas vraiment de controverse. Les théoriciens de l'intelligence semblent avoir assez unanimement supposé que le facteur g découle de causes multiples. À l'appui d'une thèse unitaire, on serait tenté de citer Spearman, qui serait pourtant un assez mauvais exemple : bien que Spearman ait été attaché à l'idée d'une *mental energy* comme origine unitaire possible du facteur g , cette position coexistait avec l'hypothèse d'une division en deux déterminants, l'aptitude à éduire (induire) et l'aptitude à reproduire (voir Spearman, 1927), qui préfigurent l'intelligence fluide et l'intelligence cristallisée proposées par Cattell (1963) ; et ses travaux ont également mentionné le rôle d'autres processus tels que la mémoire ou l'attention (Spearman, 1927). De fervents défenseurs de l'étude du facteur général pour lui-même, comme Jensen (1998) et Gottfredson (2016), s'accordent à dire que le facteur général est certainement produit par un ensemble de déterminants. Kranzler et Jensen (1991) ont montré que le facteur g peut être prédit par un ensemble de processus indépendants (voir aussi Carroll, 1991 ; Detterman, 2000) ; Gottfredson écrit qu'à sa connaissance, aucun théoricien du facteur g n'a jamais proposé qu'il était causé par un unique processus ou une unique aire cérébrale.



Cette conclusion est en réalité assez logique. La complexité de l'intelligence, et en particulier de l'intelligence fluide (l'aptitude à raisonner logiquement de façon à résoudre des

problèmes nouveaux ; voir McGrew, 2009), la place de facto au sommet de la hiérarchie de la complexité dans la cognition humaine. Il serait donc surprenant que la performance sur les tests d'intelligence ne soit pas multi-déterminée, et fasse appel à une seule aptitude parfaitement unitaire. On retrouve ici une nuance qui est loin d'être toujours explicite dans la littérature : on peut s'interroger sur le caractère uni-déterminé ou multi-déterminé du facteur général (de la *corrélation* entre les tâches), mais il est évident que la réussite à une tâche d'intelligence (la *performance* sur une tâche) est due à l'opération de plusieurs mécanismes et processus agissant de concert. Si la performance repose de toute façon sur plusieurs mécanismes, pourquoi supposer qu'un seul d'entre eux organise les différences inter-individuelles et génère la corrélation entre les tests ? L'argument de la parcimonie est parfois avancé : il serait plus parcimonieux d'expliquer le facteur général par une seule cause que par plusieurs (Carroll, 1991). Mais cet argument implique de considérer l'existence d'une entité supplémentaire "intelligence" qui cause le facteur général et qui ne se réduit pas à des processus simples et bien identifiés par ailleurs, ce qui est moins parcimonieux qu'une explication qui s'appuie uniquement sur ces processus simples (Detterman, 1982) ; en outre, cela revient à supposer qu'il n'existe aucun rôle d'autres déterminants bien connus de la performance comme la motivation, que ces déterminants n'agissent pas sur tous les tests, ou bien qu'ils ne sont que le reflet de "l'intelligence", ce qui n'est pas très parcimonieux non plus.

Il n'y a donc pas tellement, parmi les différentalistes, de tenants de l'uni-détermination des scores d'intelligence, ou même de l'uni-détermination du facteur général. Il y a en revanche au moins quatre grandes façons de concevoir un facteur *g* multi-déterminé.



1) La première est la proposition de Thomson (1916 ; voir aussi Bartholomew et al., 2009), la *sampling theory* : la performance aux tests d'intelligence est causée par un très grand nombre de déterminants ; certains de ces déterminants sont communs à tous les tests, ce qui crée la corrélation entre eux. La théorie peut incorporer des sous-ensembles de déterminants liés à certains types de tests de façon à créer une structure hiérarchique. Dans sa version originale, cette théorie souffre d'un manque de spécification de la nature des déterminants en jeu (Thomson les imagine sous la forme d'entités neuronales réparties de façon homogène à travers le cerveau), et peine surtout à expliquer pourquoi certains tests simples corrélaient bien avec le facteur général, ou encore pourquoi des tests qui semblent faire appel à des processus

très différents montrent quand même des corrélations élevées entre eux (van der Maas et al., 2006). Les trois propositions suivantes permettent de répondre à ces critiques.

2) La seconde solution est la *process overlap theory*, qui peut être vue comme une version plus moderne de la *sampling theory* (Kovacs & Conway, 2016) : la performance aux tests d'intelligence est causée par un très grand nombre de déterminants ; certains de ces déterminants sont communs à tous les tests, ce qui crée la corrélation entre eux ; par ailleurs, ces déterminants ne sont pas tous d'importance égale : le plus important et le plus central de ces déterminants communs est le contrôle cognitif. La corrélation générale représentée par le facteur *g* (et sa forte corrélation avec la capacité en mémoire de travail) est essentiellement due au rôle transversal du contrôle cognitif ; le chevauchement entre des processus plus spécifiques à un type de matériel explique l'existence de facteurs de plus bas niveau tels que les aptitudes visuo-spatiales ou quantitatives. La *process overlap theory* ajoute à la *sampling theory* la notion de centralité (le degré auquel un déterminant donné conditionne la performance de l'ensemble ; Detterman, 1987) : le fait que le contrôle cognitif soit plus central que d'autres déterminants répond aux principales critiques adressées à la *sampling theory*, et permet d'expliquer le phénomène de différenciation (la corrélation entre les tests est plus élevée pour les participants d'aptitudes faibles ; Blum & Holling, 2017 ; Detterman & Daniel, 1989 ; Spearman, 1927) par l'existence du goulot d'étranglement créé par un faible contrôle cognitif.

3) Une troisième classe de propositions théoriques considère que les déterminants de la performance sont dépendants les uns des autres : la corrélation entre les tests d'intelligence émerge non seulement parce qu'ils impliquent des déterminants communs de la performance, mais aussi parce que ces déterminants s'influencent entre eux. On retrouve cette idée dans le modèle du facteur général par mutualisme (van der Maas et al., 2006), qui propose que les différents processus cognitifs mis en jeu par les tests d'intelligence (mémoire à court terme, vitesse de traitement...) s'influencent mutuellement au cours du développement, la croissance d'un mécanisme bénéficiant à tous les autres ; ce qui conduit ces déterminants à être corrélés entre eux chez l'adulte même s'ils ne l'étaient pas à l'origine. Cette solution explique les corrélations entre facteur général et tests simples, ainsi que les corrélations entre tests de structures très différentes, par des corrélations entre les déterminants qui les influencent.

4) Une quatrième façon de considérer l'inter-dépendance entre les mécanismes contribuant au score d'intelligence est de considérer que l'intelligence correspond à un système

de processus fortement intégrés : autrement dit, la performance du système dans son ensemble dépend de chacun des processus ; une atteinte de n'importe quel processus peut amener une chute de performance massive. Cette idée est au cœur du modèle systémique de Detterman (1987 ; Detterman et al., 2016). Ce modèle met en avant l'idée de centralité des processus et d'intégration entre eux ; on y retrouve l'idée du modèle par mutualisme selon laquelle les déterminants ne sont pas indépendants les uns des autres, mais cette dépendance réside moins dans le fait que les déterminants s'influencent mutuellement que dans le fait qu'ils interagissent systématiquement pour produire la performance, l'effet de l'un étant modulé par l'efficacité de l'autre. Comme la *process overlap theory*, le modèle systémique de Detterman propose des hypothèses sur la nature des déterminants qui contribuent à l'intelligence ; contrairement à la *process overlap theory*, il ne place pas particulièrement l'accent sur le contrôle cognitif mais propose plutôt de considérer le système cognitif dans son ensemble (ce qui n'empêche pas de supposer que certains composants puissent être plus centraux que d'autres). Le facteur *g* est alors causé par le fait que tous les tests d'intelligence sont traités par le même système cognitif. L'existence d'une hiérarchie d'aptitudes regroupées par domaines peut s'expliquer facilement si l'on ajoute au modèle l'hypothèse selon laquelle le poids relatif des différents composants du modèle peut varier d'une tâche à l'autre. De récents travaux en neuroimagerie s'orientent dans une direction assez similaire à cette approche systémique en considérant qu'un test d'intelligence recrute de multiples réseaux de neurones qui se chevauchent partiellement d'une tâche à l'autre (Hampshire et al., 2012).



Ces quatre conceptions ont en commun d'adopter une position constructiviste quant à l'origine du facteur général, et de reconnaître son caractère multi-déterminé ; elles diffèrent dans la façon d'implémenter cette multi-détermination et sur les conséquences que l'on peut en tirer, mais sont en réalité largement compatibles. Je crois que la proposition systémique de Detterman est la formulation la plus intéressante, car elle est avant tout une théorie sur l'origine des *scores* d'intelligence. Les trois autres solutions théoriques sont plutôt des théories sur l'origine du *facteur général* de l'intelligence : elles cherchent à expliquer les corrélations entre tâches mais s'intéressent peu à la nature des déterminants en jeu (*sampling theory*, mutualisme), ou bien assignent un poids si important à un déterminant en particulier (*process overlap theory*) que l'on retombe sur l'hypothèse d'une cause unique ou presque unique à la corrélation entre tests (Kan et al., 2016; voir aussi Ackerman, 2016). Par contraste, l'approche systémique permet de mieux imaginer la multiplicité des déterminants mis en jeu, dans la



mesure où elle permet d'établir assez aisément un lien avec les modèles du traitement de l'information développés en psychologie cognitive.

La formalisation d'une approche systémique de l'intelligence est due à Detterman (quoiqu'on en trouve des prémices dans les propositions de Spearman, Thorndike et Thomson), mais plusieurs modèles de ce type ont été proposés. Ce ne sont pas les plus connus des modèles de l'intelligence mais ils illustrent suffisamment bien l'approche constructiviste pour mériter que l'on s'y attarde. Les principaux exemples en sont présentés à la page suivante. Dans l'ordre chronologique :

(1) La conception originelle de Detterman (1982), développée dans le contexte de travaux portant sur les déficits mnésiques chez des patients atteints de déficience intellectuelle.

(2) Une tentative de pont entre approches factorielle, causale et développementale de l'intelligence, proposée par Horn (1987). Cette proposition est particulièrement intéressante en cela qu'elle suggère que les facteurs de second ordre identifiés dans le cadre du modèle CHC représentent autant de mécanismes impliqués dans les tests d'intelligence.

(3) Une proposition de Kyllonen (2002), qui place la mémoire de travail au centre du système, dans la logique de travaux précédents insistant sur sa corrélation très élevée avec l'intelligence (Kyllonen & Christal, 1990). Les chiffres de 1 à 7 représentent des sources possibles de variabilité inter-individuelle dans le modèle ; on remarque qu'un même composant peut générer plusieurs types de variabilité (la mémoire déclarative par exemple peut varier en termes d'étendue, mais aussi en termes de vitesse d'apprentissage).

(4) Le modèle de Schneider et McGrew (2012), qui raffine la proposition de Horn en se plaçant également dans le cadre du modèle CHC. Cette conception donne un rôle un peu plus important à l'environnement que les précédentes.

(5) Le modèle de Spanoudis et Demetriou (2020), qui au premier abord diffère des précédents dans la mesure où il réintègre explicitement un déterminant appelé "facteur *g*" au centre du modèle ; mais ce facteur *g* représente en fait la somme de trois sous-systèmes liés respectivement au contrôle cognitif, à l'abstraction, et à la métacognition. Ces trois sous-systèmes centraux produisent la performance sur un test d'intelligence en association avec des processus domaine-spécifiques.

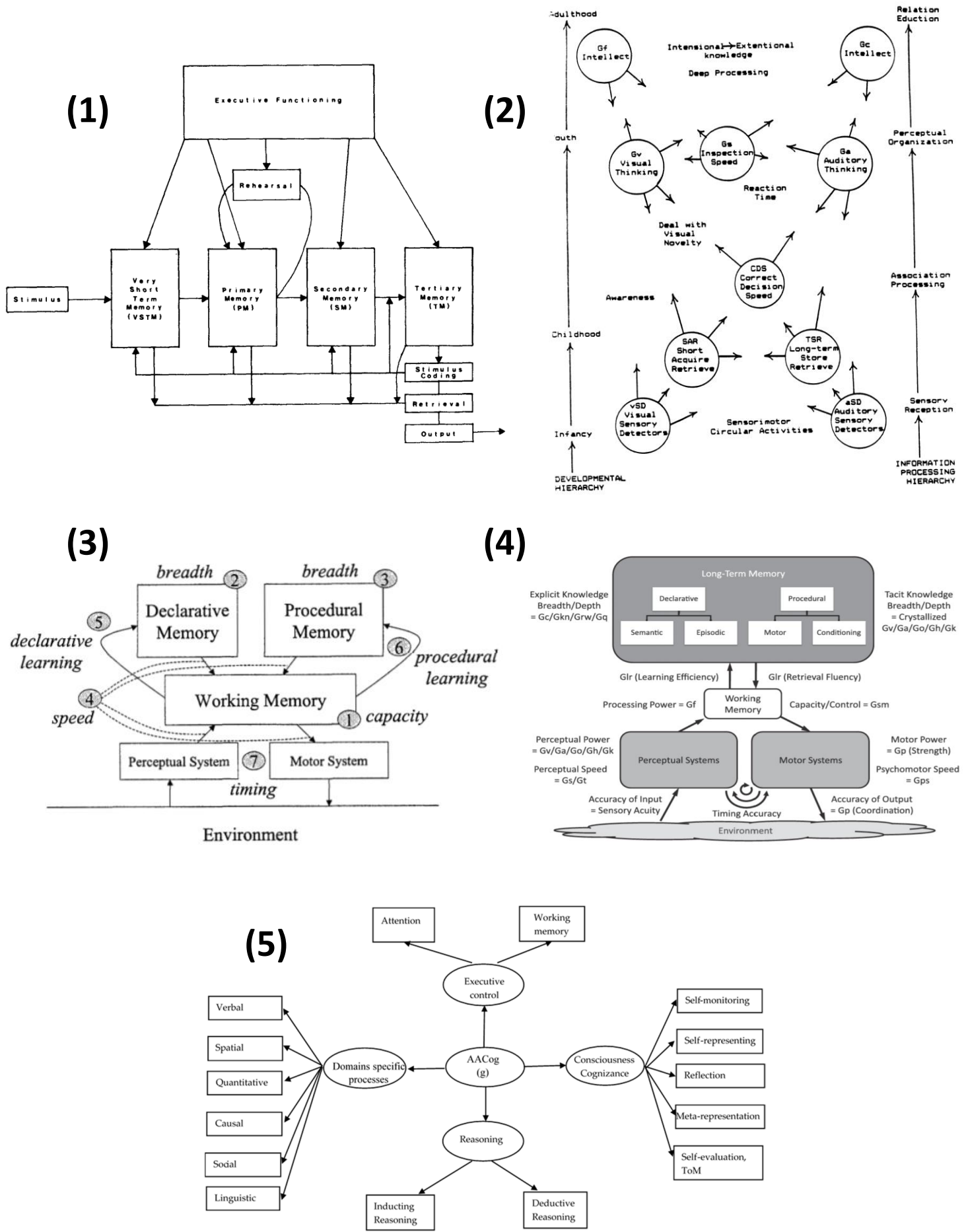


Figure 10. Cinq modèles de l'intelligence vue comme émergent d'un système complexe.

Un trait commun aux quatre premiers modèles (et discutable dans le cas du cinquième) est l'absence d'entité unitaire créant la performance d'intelligence : l'intelligence est bel et bien décrite comme une propriété émergente d'un système complexe - elle n'existe pas en tant que telle dans l'architecture cognitive et le facteur général n'est pas créé par l'opération d'un mécanisme en particulier, l'ensemble des mécanismes contribuant à la corrélation entre les tests dans la mesure où ils sont générateurs de variabilité inter-individuelle.

L'approche systémique offre une définition élégante de l'intelligence, sous la forme de "l'efficacité du système permettant le traitement de l'information dans son ensemble" ; et du facteur général, sous la forme du "degré auquel la variabilité de la performance sur différentes tâches dépend de la variabilité dans l'efficacité de celles des composantes de ce système cognitif qui contribuent à l'ensemble de ces tâches". Elle va logiquement de pair avec les idées de multi-détermination du score (de nombreux mécanismes y contribuent), mais aussi avec l'idée de multi-réalisabilité (on peut imaginer plusieurs combinaisons de mécanismes menant au même score, un déficit en mémoire procédurale pouvant par exemple se compenser par une excellente vitesse perceptive). Elle met aussi l'accent sur le caractère intégré des déterminants de l'intelligence, d'une façon qui invite à réfléchir à leurs interactions et à leurs liens de dépendance possibles.

L'un des intérêts de cette idée est qu'en formalisant la notion de goulot d'étranglement, elle permet de mieux concevoir comment une chute brutale peut se produire sur une classe de tâches particulière lorsqu'un déterminant nécessaire à ce type de tâches est affecté : un déficit perceptif ou moteur (à l'entrée ou à la sortie du système) pour une modalité sensorielle particulière génère logiquement un déficit massif pour cette modalité, ce qui donne un sens à des confusions fréquentes en clinique - par exemple entre dyspraxie et déficience intellectuelle. Cette idée invite en retour à considérer que le poids des différents déterminants du système varie d'un individu à l'autre ou d'un groupe à l'autre : le poids d'un mécanisme central sur la performance sera plus important s'il est fortement déficitaire - la corrélation entre acuité visuelle et performance aux matrices de Raven sera nulle chez l'étudiant en psychologie moyen, élevée chez des participants malvoyants.

Il s'agit d'une limite importante des résultats obtenus dans le cadre d'analyses factorielles classiques, presque toujours réalisées sur le même type d'échantillon sain ; cette limite matérialise la remise en cause du postulat d'homogénéité locale évoqué plus haut

(page 9), et illustre les précautions nécessaires lorsque l'on souhaite interpréter les résultats d'un test d'intelligence dans un contexte clinique, le facteur général n'étant pas forcément un aussi bon résumé des performances d'un participant déficitaire (auquel le test se destine pourtant !) que d'un participant sain. C'est plus généralement une limite des modèles qui postulent une cause unique au facteur général, ou testent la corrélation entre facteur général et un ou plusieurs mécanismes : le déterminant qui a le poids le plus important en moyenne n'est pas forcément celui qui a le poids le plus important pour un individu donné. Le contrôle cognitif joue un rôle central selon la *process overlap theory*, mais cette théorie s'appuie sur des études qui n'ont jamais cherché à inclure des participants atteints de déficience intellectuelle, pour lesquels le mécanisme qui contraint le plus fortement le score peut être autre. Dans une approche systémique, la recherche de "la cause de l'intelligence" est par définition vouée à l'échec.



On retrouve ici la nuance entre "origine du facteur général" et "origine de la performance d'intelligence", ou entre "entité statistique" et "entité causale" : un mécanisme central, critique pour le fonctionnement du système, peut ne générer aucune variabilité inter-individuelle en l'absence de pathologie, et donc rester invisible à l'analyse factorielle et ne pas corrélérer avec le facteur général (Jensen, 1987), tout en étant susceptible de générer un effondrement de la performance lorsqu'il est altéré. Ce principe est le pendant de la remarque précédente selon laquelle une entité statistique comme le facteur général peut correspondre à plusieurs causes : de façon symétrique, un mécanisme causal peut ne pas générer de variance et donc ne pas avoir d'effet statistique en moyenne. Là encore, le niveau de description inter-individuel ne résume pas le niveau de description individuel. Un lien entre les deux niveaux est possible, comme l'illustrent les tentatives de traduire le modèle CHC sous la forme d'un système cognitif (Horn, 1987 ; Schneider & McGrew, 2012), mais il n'existe aucune garantie que cette traduction soit fidèle puisqu'il n'existe pas d'identité entre la variance sur une tâche et la variance sur un processus sous-jacent. (Cette même critique s'applique aux modèles en *blended variables* : e.g. Ashton et al., 2009.)



Ce problème de fond constitue une réponse à l'argument souvent employé par les défenseurs de *g* selon lequel "l'origine du facteur général n'a pas d'importance", ou bien "on constate que le facteur *g* joue bien son rôle de compétence agrégée et il peut donc être vu comme ayant directement un rôle causal" (e.g. Gottfredson, 1997b), ou encore "même si

l'intelligence est d'origine multiple à un certain niveau de description elle devient unitaire au niveau comportemental" : autrement dit, l'argument selon lequel un score d'intelligence total est un résumé fidèle des aptitudes agrégées de la personne, quels que soient les facteurs qui en sont à l'origine. C'est peut-être vrai en moyenne et dans le cadre des échantillons sur lesquels les analyses factorielles sont réalisées traditionnellement, mais certainement pas lorsque le postulat d'homogénéité locale n'est pas vérifié, c'est-à-dire lorsque la contribution relative des déterminants de la performance est susceptible d'être différente de la moyenne. L'interprétation située au niveau des mécanismes individuels devient alors critique pour l'interprétation des déficits observés en contexte clinique, mais aussi pour les comparaisons entre groupes de participants hétérogènes.



2. Le rôle des stratégies dans les tests d'intelligence

Parmi les mécanismes cognitifs de la performance sur les tests d'intelligence, les stratégies sont peut-être celui qui a reçu le moins d'attention. Il existe une littérature sur les comportements stratégiques dans certains domaines cognitifs de haut niveau, notamment la cognition numérique (voir par exemple Lemaire & Reder, 1999 ; Lemaire & Siegler, 1995), mais beaucoup moins dans les tâches d'intelligence (fluide) proprement dites. On retrouve essentiellement trois ensembles de travaux dans la littérature.



Le premier porte sur les cubes de Kohs et montre qu'il existe trois stratégies : globale, analytique et synthétique (Rozenchwajg, 1991 ; Rozenchwajg & Corroyer, 1995 ; Rozenchwajg et al., 2005, 2009). Les stratégies analytique et synthétique sont plus efficaces, et sont plus utilisées par les participants de performance élevée (les jeunes adultes) que par les participants de performance faible (les enfants, les personnes âgées, les personnes atteintes de déficience intellectuelle). Le second ensemble de travaux porte sur les tâches de syllogismes et autres résolutions de problèmes à trois termes, et montre qu'il existe des stratégies verbales et visuo-spatiales (Egan & Grimes-Farrow, 1982 ; Ford, 1995) ; les participants atteints de dyslexie ont par exemple plus tendance à utiliser une stratégie visuo-spatiale plutôt que verbale (Bacon et al., 2007 ; Bacon & Handley, 2010). Le troisième porte sur les tâches de raisonnement visuo-spatiales à choix multiple, telles que le *paper folding test*, les tâches d'analogies ou les matrices de Raven, et distingue deux stratégies : le *constructive matching*, stratégie efficace consistant à reconstruire mentalement la réponse manquante, et l'élimination de réponses, stratégie moins efficace consistant à éliminer les réponses possibles qui ne semblent pas correspondre (Snow, 1978, 1980 ; Bethell-Fox et al., 1984).

Les stratégies de constructive matching et d'élimination de réponses ont fait l'objet d'étonnamment peu de recherches, au vu de leur transversalité et du rôle qu'elles sont supposées jouer dans les matrices de Raven - de loin la tâche d'intelligence fluide la plus utilisée en recherche (Ackerman et al., 2005), et généralement considérée comme le meilleur indicateur de l'intelligence fluide (Carpenter et al., 1990). On sait que le constructive matching va de pair avec un temps de réponse plus élevé, une plus grande confiance des participants dans leur réponse, une performance plus élevée, et une probabilité moindre d'être influencé par les plus réponses possibles les plus saillantes (Arendasy & Sommer, 2013; Becker, Schmitz, et al., 2016 ; Mitchum & Kelley, 2010; Vigneau et al., 2006), mais guère plus.

L'existence de ces deux stratégies apparaît pourtant comme particulièrement intéressante au regard de la corrélation observée entre capacité en mémoire de travail et performance sur les tâches d'intelligence fluide. Comme évoqué dans le chapitre précédent (page 65 et page 95), cette corrélation entre mémoire de travail et intelligence est au cœur d'une bonne partie des travaux contemporains portant sur la mémoire de travail, et de plusieurs des modèles cherchant à expliquer les mécanismes de la performance sur les tests d'intelligence (Kyllonen, 2002 ; Kovacs & Conway, 2016) ; mais elle est toujours relativement mal comprise. J'ai déjà évoqué la possibilité selon laquelle la valeur prédictive de la capacité en mémoire de travail pourrait s'expliquer par un lien avec l'usage de stratégies efficaces : les participants avec une forte capacité en mémoire de travail auraient plus tendance à utiliser une stratégie plus efficace dans la tâche. L'existence de deux stratégies bien identifiées sur les matrices de Raven, qui montrent une relation stable avec la capacité en mémoire de travail (Ackerman et al., 2005), offrait donc une excellente occasion de tester cette hypothèse, et de mettre en lumière l'importance de bien comprendre les stratégies dans les tests d'intelligence.



J'ai donc cherché à tester le lien entre mémoire de travail et usage du constructive matching sur les matrices de Raven (Gonthier & Thomassin, 2015) dans une étude composée de deux expériences. Dans la première expérience, nous avons évalué le degré auquel chaque participant avait utilisé le constructive matching et l'élimination de réponses, à l'aide de 4 questions en auto-report présentées en fin de tâche (par exemple : *J'ai pris le temps de réfléchir à la bonne réponse avant de regarder les réponses proposées*). Nous avons observé une corrélation positive entre capacité en mémoire de travail et utilisation du constructive matching, et une corrélation négative entre capacité en mémoire de travail et utilisation de l'élimination de réponses ; mais surtout, nous avons trouvé une médiation totale de la corrélation entre capacité en mémoire de travail et performance sur les matrices de Raven par les deux indicateurs stratégiques (médiation multiple par bootstrap ; Preacher & Hayes, 2008).

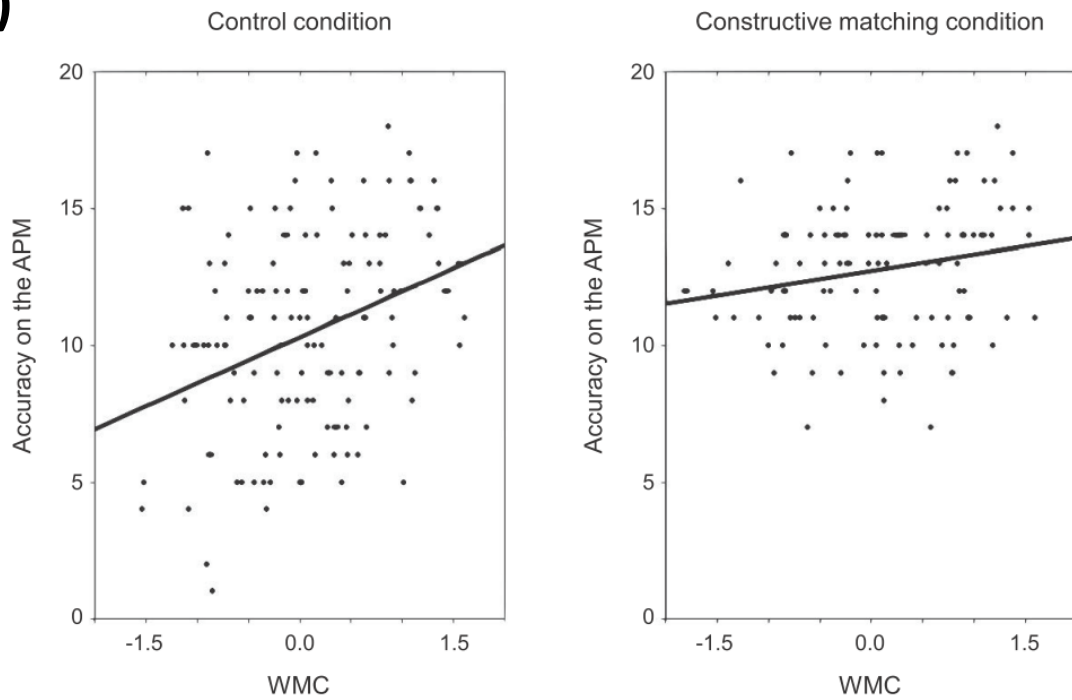
Dans la seconde expérience, nous avons cherché à vérifier le rôle causal des stratégies en les manipulant directement grâce à une approche expérimentale-corrélationnelle (Cronbach, 1957 ; voir aussi le chapitre précédent, page 99). Pour ce faire, nous avons créé une condition visant à induire l'usage du constructive matching chez les participants, en présentant la matrice par morceaux (un protocole inspiré de Luria & Tsvetkova, 1964). Le mode de représentation reproduisait les mouvements oculaires de participants cherchant à reconstruire



la bonne réponse (Carpenter et al., 1990) : nous présentions d'abord la première ligne de la matrice, puis la seconde, puis la troisième, et enfin les différentes réponses possibles. Les participants étaient libres de passer à l'étape suivante quand ils le souhaitaient ; cet étayage ne les forçait donc pas à passer plus de temps sur l'item et ne donnait aucune indication supplémentaire sur la bonne réponse. Nous avons observé que la corrélation entre capacité en mémoire de travail et performance sur les matrices diminuait significativement dans la condition visant à induire un constructive matching, par rapport à la condition contrôle. Les résultats de ce travail sont résumés dans la Figure 11.

Les résultats de cette étude suggèrent donc que l'usage de stratégies efficaces joue un rôle causal dans la corrélation entre capacité en mémoire de travail et performance sur les tests d'intelligence : le lien entre mémoire de travail et performance n'est plus significatif lorsque l'on contrôle pour l'usage de stratégies efficaces, et la relation semble également disparaître lorsque l'on encourage les participants à utiliser la stratégie efficace de constructive matching. Nos conclusions ont été contestées avec véhémence par les tenants d'une explication endogène, et plus particulièrement par Jastrzębski et collaborateurs (2018), qui ont échoué à répliquer les résultats de notre analyse de médiation dans une étude utilisant un protocole similaire. Une partie de ces critiques naît de l'incompétence malheureuse des auteurs (qui, entre autres erreurs, ignorent la possibilité de tester l'interaction entre une variable catégorielle et une variable continue, s'étonnent qu'un score composite calculé en moyennant trois scores standardisés corrélés entre eux ait un écart-type inférieur à 1, et confondent corrélation et corrélation corrigée pour atténuation) ; on peut aussi s'interroger sur la pertinence de tester le rôle des stratégies dans une version des matrices de Raven en temps limité, ce qui a de fortes chances de restreindre l'usage du constructive matching (e.g. Friedman & Miyake, 2004 ; St-Clair Thompson, 2007). Mais la non-réplication subsiste et interroge sur la généralisabilité de nos résultats.

(1)



(2)

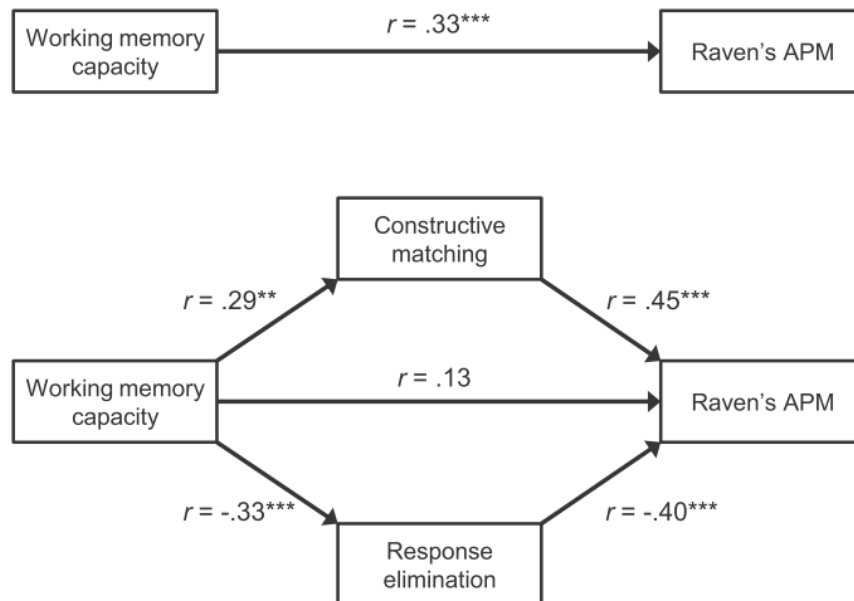


Figure 11. (1) Corrélations entre capacité en mémoire de travail, comportements stratégiques et performance sur les matrices de Raven. (2) Corrélations entre mémoire de travail et performance sur les matrices de Raven dans une condition contrôlée (à gauche) ou induisant un constructive matching (à droite). Tiré de Gonthier, C., & Thomassin, N. (2015). Strategy use fully mediates the relationship between working memory capacity and performance on Raven's matrices. *Journal of Experimental Psychology: General*, 144(5), 916-924.

L'étude précédente présentait une limite que j'ai souhaité approfondir : nous n'avions évalué le fonctionnement stratégique qu'au niveau global de la tâche, sous la forme d'un score total. Mais on peut également envisager que le déploiement de stratégies varie d'un item à l'autre. La variabilité intra-individuelle dans l'utilisation des stratégies est l'une des quatre principales dimensions de la variabilité stratégique (avec les différences dans la composition du répertoire stratégique, dans la fréquence d'usage relative des différentes stratégies, et dans l'efficacité d'utilisation des différentes stratégies ; voir Lemaire & Siegler, 1995). En outre, cette variation intra-individuelle n'est pas forcément aléatoire : elle peut refléter une adaptation aux contraintes de la tâche, par exemple sous la forme d'un satisficing qui engendre un repli vers des stratégies plus simples lorsque la tâche devient trop complexe (e.g. Klayman, 1985 ; Payne, 1976), et elle est parfois prédite par des variables inter-individuelles telles que la capacité en mémoire de travail (Reder & Schunn, 1999 ; Schunn & Reder, 1998, 2001). Dans le cas particulier des matrices de Raven, plusieurs auteurs ont suggéré que l'usage du constructive matching pouvait décliner au fil de la tâche (Bethell-Fox et al., 1984 ; Snow, 1978, 1980), sans l'avoir testé directement de façon systématique.



Dans une seconde étude, je me suis donc intéressé à la variabilité intra-individuelle des stratégies de constructive matching et d'élimination de réponses sur les matrices de Raven, et à deux prédicteurs possibles de cette variabilité intra-individuelle (Gonthier, 2017b, 2018 ; Gonthier & Roulin, 2019). Le premier de ces prédicteurs était la capacité en mémoire de travail, qui devrait être d'autant plus prédictive du constructive matching que le nombre de règles logiques qui constituent les items augmente : l'intégration d'un très petit nombre de contraintes pour reconstituer la réponse manquante ne devrait pas imposer une charge importante en mémoire (Mulholland et al., 1980). Le second était le besoin de cognition, c'est-à-dire la tendance à s'engager dans une réflexion approfondie et à y prendre plaisir, un trait motivationnel qui prédit la réussite aux matrices de Raven (Hill et al., 2013, 2016) et devrait aussi être d'autant plus prédictif du constructive matching que les items deviennent plus complexes et que le coût conceptuel de cette stratégie augmente.



Les résultats sont illustrés dans la Figure 12. Ils ont permis de confirmer les hypothèses de la littérature concernant l'existence d'un déclin progressif du constructive matching et d'une augmentation progressive de l'élimination de réponses au fil de la tâche, indiquant que les participants s'orientent progressivement vers la stratégie moins efficace lorsque le coût

conceptuel de la stratégie plus efficace augmente. Les résultats indiquent par ailleurs que cette variation intra-individuelle dépend de la capacité en mémoire de travail : les participants maintiennent l'usage du constructive matching d'autant plus longtemps que leur capacité en mémoire de travail est élevée. Le besoin de cognition prédit l'utilisation du constructive matching, mais non la variation intra-individuelle - un besoin de cognition élevé va de pair avec une utilisation plus importante du constructive matching à travers l'ensemble de la tâche. Les seuls participants qui continuent à utiliser le constructive matching sur les derniers items des matrices sont ceux qui présentent à la fois une forte capacité en mémoire de travail et un fort besoin de cognition.

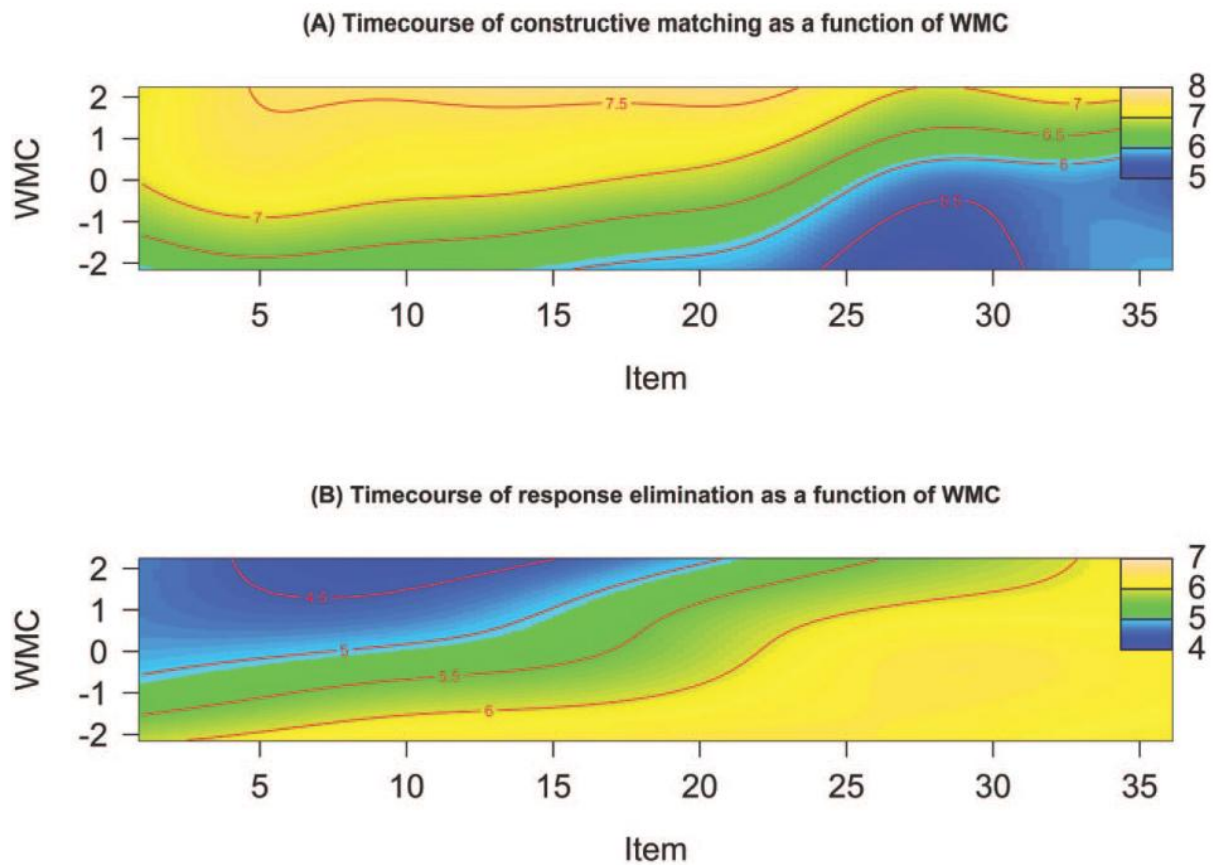


Figure 12. Variation intra-individuelle dans l'usage du constructive matching (A) et de l'élimination de réponses (B) dans les matrices de Raven, en fonction de la capacité en mémoire de travail (axe des ordonnées). Tiré de Gonthier, C., & Roulin, J.-L. (2019). Intraindividual strategy shifts in Raven's matrices, and their dependence on working memory capacity and need for cognition. *Journal of Experimental Psychology: General*, 149(3), 564-579.

L'utilisation du constructive matching apparaît donc conditionnée à la fois par une dimension cognitive, la mémoire de travail, et par une dimension motivationnelle, le besoin de cognition. En retour, l'impact de ces deux variables sur le type qualitatif de stratégie utilisée peut être un élément pour comprendre pourquoi elles corrèlent avec la performance sur la tâche - autrement dit, comment un niveau élevé de mémoire de travail ou de besoin de cognition peut se traduire concrètement par une réussite plus importante. Cet exemple illustre bien, à mon sens, plusieurs thématiques récurrentes dans ce document de synthèse : l'intérêt d'examiner les mécanismes qualitatifs pour comprendre comment une "aptitude élevée" donne lieu à un score élevé, mais aussi l'intérêt de considérer plusieurs déterminants possibles pour mieux appréhender la complexité de la performance, le système mis en jeu dans la tâche étant constitué de mécanismes cognitifs, mais aussi motivationnels, situationnels, etc. L'existence d'une variation intra-individuelle qui interagit de façon complexe avec ces déterminants souligne en outre les limites d'un raisonnement unidimensionnel purement inter-individuel qui aboutit à calculer uniquement un score total.



Mon principal sujet d'insatisfaction concernant ces travaux portait sur la mesure des stratégies : dans les deux cas, j'ai utilisé une mesure en auto-report, constituée de quatre items présentés à la fin de la tâche pour la première étude, et de deux items présentés après chaque item pour la seconde. De façon générale, je suis fortement en désaccord avec le rejet arbitraire des reports verbaux, souvent considérés comme insuffisamment sérieux ou pas assez raffinés par les chercheurs en psychologie cognitive ; l'exemple des processus stratégiques utilisés en mémoire visuo-spatiale montre bien que toutes les subtilités comportementales ne sont pas accessibles aux méthodes objectives, mais indirectes, telles que l'analyse des temps de réponse ou des mouvements oculaires (Gonthier, 2020 ; voir page 93). Les reports verbaux ont le mérite d'être directs, relativement exhaustifs, et constituent une source de données légitime si l'on prend soin de réfléchir aux biais potentiels (Ericsson & Simon, 1980). Mais dans ce cas particulier, ils montrent leur limite. D'une part, ils ne sont pas utilisables pour des échantillons de jeunes enfants ou de participants atteints de déficience intellectuelle, qui sont pourtant des terrains d'étude très intéressants pour le rôle causal des stratégies dans la performance (voir page 125). D'autre part, le faible nombre de questions que l'on peut imaginer pour interroger des participants sur l'utilisation de deux stratégies limite nécessairement la fidélité de la mesure (c'était l'une des principales critiques de Jastrzębski et al., 2018, pour le coup parfaitement fondée).

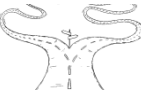
Même s'il existait une infinité de façons possibles de demander à un participant s'il a ou non passé du temps à regarder la matrice avant de répondre, évaluer la variabilité intra-individuelle contraint fortement le nombre de questions que l'on peut poser après chaque item sans interférer avec la passation. Cette question s'est posée dans le cadre de l'étude précédente, où le constructive matching et l'élimination de réponses n'étaient évalués que par une question chacune après chaque item. Il est difficile d'évaluer la fidélité d'un item unique présenté après chaque item (calculer une fidélité totale sur la tâche n'aurait évidemment pas de sens dans le contexte où c'est la précision de la mesure au niveau de chaque item qui importe). Dans l'étude précédente, j'ai utilisé une solution alternative consistant à estimer la fidélité en split-half de la trajectoire intra-individuelle (en calculant la pente de l'effet aléatoire représentant cette trajectoire à partir des items pairs, puis des items impairs, et enfin en calculant la corrélation entre ces deux pentes avec correction de Spearman-Brown). Cette solution partiellement satisfaisante a montré une fidélité limitée et légèrement inférieure à .70.

Nous avons donc cherché à développer une meilleure façon d'évaluer le fonctionnement stratégique sur les matrices de Raven. La principale alternative à la méthode de l'auto-report est l'étude des mouvements oculaires, qui permettent d'examiner le temps passé à regarder la matrice (plutôt que les réponses possibles), le délai passé à regarder la matrice avant d'aller pour la première fois examiner les réponses possibles, et le nombre d'alternances entre la matrice et les réponses possibles : trois indicateurs qui reflètent le fonctionnement supposé des deux stratégies (Bethell-Fox et al., 1984 ; Vigneau et al., 2006). Ces trois mesures peuvent être calculées en proportion du temps passé sur l'item pour contrôler les différences de rapidité, entre groupes d'âges par exemple. Cette méthodologie donne de bons résultats mais est difficile à déployer à grande échelle, ou dans des échantillons difficiles d'accès (établissements scolaires...).



Nous avons récemment travaillé sur une méthode alternative, qui consiste à reproduire le fonctionnement d'une analyse des mouvements oculaires en modifiant la structure des matrices de Raven (Rivollier, Quinton, Gonthier, & Smeding, 2020). Le principe est tout simplement de décomposer la présentation de chaque matrice en deux parties : on affiche soit la moitié supérieure de l'item (la matrice dont il faut trouver l'élément manquant), soit la moitié inférieure (les différentes réponses proposées). Le participant peut cliquer librement dans la moitié inférieure ou supérieure pour passer de l'un à l'autre. On peut ainsi reconstituer

les trois mêmes indicateurs de la stratégie utilisée : temps passé à regarder la matrice, latence avant d'aller regarder pour la première fois les réponses proposées, et nombre d'alternances entre matrices et réponses proposées. (Ce dispositif ressemble fortement à la méthode proposée par Corroyer & Rozencwajg, 1995; Rozencwajg & Corroyer, 2001, pour l'étude des cubes de Kohs avec la tâche SAMUEL, qui demande de cliquer sur un bouton à chaque fois que le participant souhaite revoir le modèle à reproduire.)



Plusieurs expériences utilisant cette méthode (Rivollier, Quinton, Gonthier, & Smeding, 2020) nous ont permis de vérifier que cette présentation découpée des matrices de Raven n'altérerait pas significativement les propriétés psychométriques de la tâche, ni la performance moyenne des participants ; et que les trois indices stratégiques qu'elle permet de calculer corrèlent avec la performance totale dans le sens attendu si ces indices reflètent effectivement l'usage des deux stratégies de constructive matching et d'élimination de réponses. La seule différence majeure avec l'enregistrement des mouvements oculaires réside dans le nombre d'alternances entre matrices et réponses possibles, qui est beaucoup plus faible avec notre procédure ; le fait que les participants doivent déplacer la souris et cliquer pour afficher l'autre moitié de l'item contraint certainement le nombre d'alternances qu'ils réalisent. (Une autre expérience nous a permis de vérifier, au passage, que ce dispositif permet une induction du constructive matching proche dans la logique de celle que j'avais utilisée précédemment dans Gonthier & Thomassin 2015, mais moins ambiguë, en imposant tout simplement un délai minimal avant de permettre aux participants de consulter pour la première fois les réponses possibles.) Cette nouvelle méthode donne donc accès à une analyse des stratégies utilisées par les participants au niveau de chaque item, sans utiliser la méthode de l'auto-report. Il serait utile de vérifier dans quelle mesure ces résultats convergent avec les indicateurs obtenus en auto-report, mais ce dispositif me paraît déjà prometteur en l'état ; ce qui m'a conduit à l'utiliser dans des projets de recherche présentés plus loin (pages 152-153).



3. Le rôle du temps dans les tests d'intelligence

Le temps de réponse joue un rôle assez ambigu dans les tests d'intelligence. On peut le considérer comme un aspect de la performance : un participant qui répond rapidement aux problèmes qui lui sont posés est un participant efficace, donc "intelligent". La vitesse est vue comme valorisante. Les tests d'aptitude centrés sur la vitesse (le test des dominos en est un exemple classique) demandent une résolution rapide des problèmes de façon à laisser le temps au participant de traiter suffisamment d'items pour atteindre un score élevé. Par ailleurs, la vitesse de traitement corrèle positivement avec la performance aux tests d'intelligence (e.g. Jensen, 1987 ; Vernon, 1983) et constitue, dans le cadre du modèle CHC, un facteur de seconde strate (*gs*) qui représente une dimension de l'aptitude générale (McGrew, 2009). Selon cette logique, le temps passé à résoudre un problème est donc un indicateur de la réussite, au même titre qu'une bonne réponse, et ne constitue en aucun cas un *déterminant* de la performance ; une bonne performance va de pair avec un temps de réponse faible.

D'un autre point de vue, résoudre un problème demande un temps minimal de réflexion : aucun participant ne parvient à la bonne réponse sur un item complexe des matrices avancées de Raven sans investir plusieurs dizaines de secondes dans la recherche des règles logiques. L'utilisation d'une stratégie complexe mais efficace, comme le constructive matching, demande du temps : le constructive matching corrèle positivement, et l'élimination de réponses négativement, avec le temps de réponse à un problème (e.g. Gonthier & Thomassin, 2015). Dans l'étude précédente, j'ai observé que les items les plus difficiles des matrices de Raven génèrent non seulement une chute de performance, mais aussi un désengagement des participants qui se manifeste par une diminution des temps de réponse ; seuls les participants qui ont à la fois une capacité en mémoire de travail élevée et un besoin de cognition élevés continuent à présenter un temps de réponse élevé sur les items les plus difficiles (Gonthier & Roulin, 2019). Selon cette logique, un temps de réponse élevé contribue à la réussite : le temps de réponse est un déterminant de la performance ou, en tout cas, un symptôme de la mise en œuvre de processus complexes qui déterminent la performance ; une bonne performance va de pair avec un temps de réponse élevé.

Ces deux perspectives ne sont pas vraiment incompatibles. Sur les matrices de Raven, qui ont fait l'objet du plus grand nombre de recherches, plusieurs études ont conclu que les réponses données avec un temps de réponse faible ou élevé peuvent représenter des



processus différents (Partchev & De Boeck, 2012 ; DiTrapani et al., 2016). Une étude récente (Perret & Dauvier, 2018) a montré que les enfants de différents niveaux d'aptitude se distinguent notamment par la façon dont ils modulent leur temps de réponse face à la difficulté des problèmes : les réponses sont universellement rapides sur les items les plus simples, mais les enfants les plus performants sont ceux dont le temps de réponse augmente le plus sur les problèmes les plus difficiles. Le lien entre temps de réponse et performance dépend donc de la complexité des problèmes. La littérature permet de conclure que l'on observe plutôt une corrélation négative entre temps de réponse et performance pour les problèmes simples (pour lesquels un temps de réponse important témoigne d'une difficulté à induire les règles), mais une corrélation positive pour les problèmes difficiles (pour lesquels un temps de réponse important témoigne de la mise en œuvre de processus complexes de résolution de la tâche ; Goldhammer et al., 2015 ; Becker, Schmitz, Görizt, & Spinath, 2016). Ce constat, selon lequel le temps de réponse peut jouer un rôle positif dans la détermination de la performance, m'a amené à me poser deux questions différentes.

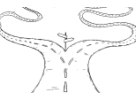
Ma première interrogation (Gonthier, en préparation) concernait l'influence de la pression temporelle dans une tâche de raisonnement. Les tests d'intelligence varient considérablement dans l'imposition ou non d'une limite de temps : beaucoup sont obligatoirement présentés en temps contraint ; d'autres tests, comme les matrices de Raven sont souvent passés en temps limité, le manuel pouvant même suggérer un temps maximum d'administration, mais sont très explicitement prévus pour être passés en temps libre (Raven, 2008). Les échelles de Wechsler ne donnent pas de limite de temps pour les subtests de matrices mais encouragent le psychologue à passer à l'item suivant si le participant met trop de temps à répondre. Quelle est l'influence de la pression temporelle sur la réussite ? Si le temps corrèle négativement avec la performance, imposer une pression temporelle devrait plutôt renforcer les différences inter-individuelles dans le sens attendu, les participants dotés d'une aptitude élevée étant capables de résoudre les items dans le temps imparti tandis que les participants d'aptitude faible échoueront. Mais si le temps corrèle positivement avec la performance, on peut s'attendre à ce que l'effet d'une pression temporelle soit plutôt contre-productif : la limite de temps devrait plutôt empêcher les participants d'aptitude élevée de trouver la bonne réponse, notamment sur les items les plus difficiles, écrasant ainsi les différences inter-individuelles. On voit que cette question a des applications cliniques directes pour la conception de tests d'intelligence.



La pression temporelle pourrait en particulier interférer avec l'utilisation de stratégies, une limite de temps empêchant l'utilisation du constructive matching par les participants de niveau élevé. La littérature fournit plusieurs arguments en ce sens, notamment les études utilisant la pression temporelle pour limiter l'usage de stratégies complexes (e.g. Friedman & Miyake, 2004 ; St-Clair Thompson, 2007) et les études qui montrent une chute de performance plus importante pour les participants d'aptitude élevée en situation de pression temporelle (e.g. Beilock & Carr, 2005). On peut d'ailleurs se demander si l'échec de Jastrzębski et collaborateurs (2018) à répliquer la médiation que nous avons montrée du lien entre capacité en mémoire de travail et score aux matrices de Raven par le comportement stratégique n'est pas dû, en partie, à leur utilisation d'une version des matrices de Raven en temps limité.

J'ai donc réalisé une étude (Gonthier, en préparation) pour examiner l'effet de la pression temporelle sur les matrices de Raven (réalisées en temps libre, avec une limite de temps de 20 minutes correspondant à la recommandation du manuel, ou avec une limite de temps de 10 minutes). Je me suis intéressé non seulement au niveau inter-individuel du score total (on peut attendre un effet trivial de diminution de la performance totale lorsqu'une limite de temps est imposée), mais surtout au niveau plus détaillé des mécanismes mis en jeu au niveau de chaque item. Outre la réussite, j'ai examiné plusieurs indicateurs pouvant nous renseigner sur le processus de résolution : temps de réponse, confiance du participant dans sa réponse, et enfin usage auto-rapporté du constructive matching et de l'élimination de réponses. Je m'attendais à observer un effet délétère de la pression temporelle en particulier pour les items les plus difficiles en fin de tâche, qui nécessitent le plus de temps pour être résolus ; je m'attendais aussi à ce que la pression temporelle fasse diminuer l'usage du constructive matching et augmenter l'usage de la stratégie moins complexe d'élimination de réponses. Enfin, je m'attendais à ce que la pression temporelle diminue la corrélation entre la performance d'une part, et la capacité en mémoire de travail et le besoin de cognition d'autre part, la limite de temps générant une chute de performance pour les participants d'aptitude élevée qui devraient donc présenter une différence moins marquée avec les participants les plus faibles.

Quelques-uns des résultats de cette étude sont présentés dans la Figure 13 à titre d'illustration. Ils se sont avérés intéressants mais assez éloignés de mes hypothèses. La pression temporelle est à l'origine d'une diminution globale de la performance, sans interaction avec la



difficulté de l'item ; on trouve d'ailleurs une chute de performance en situation de pression temporelle dès le tout premier item présenté, pourtant très simple. On retrouve l'augmentation du temps de réponse en fonction de la difficulté des items que relève la littérature ; mais la pression temporelle conduit à une plus faible modulation du temps de réponse à travers toute la tâche, et en particulier sur les items difficiles en fin de test. La pression temporelle est à l'origine d'une confiance plus faible dans la réponse donnée, et d'une utilisation plus faible du constructive matching à travers toute la tâche, sans interaction avec la difficulté de l'item. Mais cette diminution du constructive matching n'est pas compensée par une hausse de l'élimination de réponses, qui diminue également (ce qui suggère qu'il manque à la littérature une troisième stratégie, le guessing, qui doit augmenter dans ce contexte).

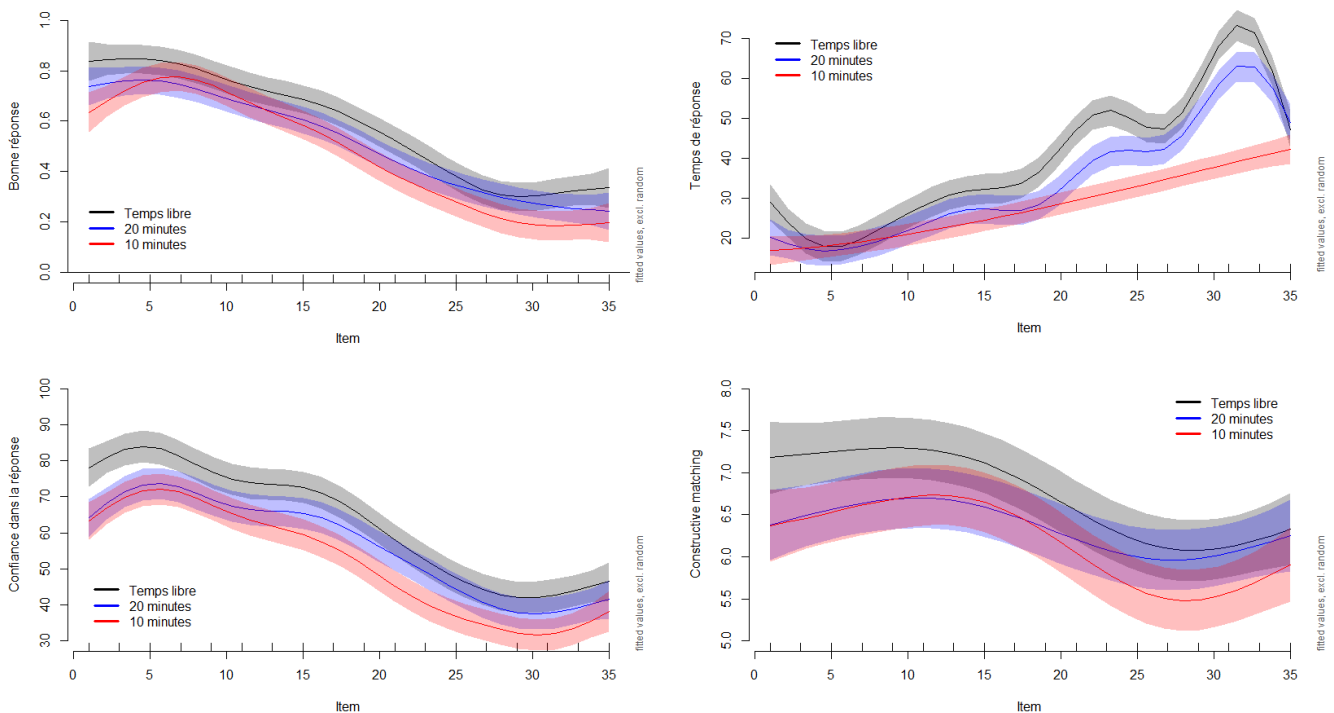


Figure 13. Effet de la pression temporelle sur la probabilité de donner une réponse, le temps de réponse (en secondes), le degré de confiance subjective du participant dans sa réponse (en %), et l'usage du constructive matching (sur une échelle de 1 à 9).

Dans l'ensemble, la pression temporelle semble conduire à une accélération des réponses, les participants prenant moins de temps pour réfléchir et contrôler les règles logiques qu'ils ont induites et pour utiliser une stratégie systématique quelle qu'elle soit. Étonnamment, les corrélations avec besoin de cognition et capacité en mémoire de travail se sont avérées insensibles à la pression temporelle, ce qui suggère que la limite de temps ne

modifie pas fondamentalement la nature de la tâche et n'affecte pas plus les participants qui ont la meilleure aptitude. Mes résultats ne remettent donc pas en cause l'usage de tests en temps limité, mais suggèrent que cette procédure sous-estime la compétence réelle d'un participant.

Nous avons mené une autre étude (Gonthier, Trégouët, Bac, & Gavornikova-Baligand, en préparation) en lien avec la question du temps, dans le cadre d'un projet visant à mieux comprendre la performance de raisonnement souvent très faible observée chez les élèves de classe SEGPA (classe adaptée destinée aux élèves qui présentent de fortes difficultés d'apprentissage). Ces élèves tendent à échouer massivement aux tâches de raisonnement logique, mais ne présentent pourtant pas de déficience intellectuelle - il s'agit d'un critère d'exclusion qui oriente vers d'autres types de classes adaptées. Dans la continuité logique des travaux portant sur l'effet de la menace du stéréotype dans les tâches cognitives de haut niveau (e.g. Huguet, Brunot, & Monteil, 2001), en particulier en ce qui concerne l'effet du sexe (Huguet & Régner, 2007), nous avons d'abord fait l'hypothèse qu'un mécanisme similaire pouvait être à l'œuvre ici. En effet, les aptitudes cognitives de ces élèves sont systématiquement dévalorisées, d'une façon qui pourrait les conduire à chuter sur ce type de tâche complexe. Mais nous avons échoué à mettre en évidence un effet de menace du stéréotype qui affecterait spécifiquement ces élèves. Nous avons en revanche observé une différence importante dans le mode de résolution de la tâche.



Nos principaux résultats sont illustrés dans la Figure 14. Comme attendu, nous avons observé une performance nettement plus faible chez les élèves de SEGPA, qui de la 6ème à la 3ème présentaient un score de raisonnement plus faible aux matrices de Raven standard que des élèves de 6ème ordinaire. Mais cette différence allait de pair avec une importante différence dans les temps de réponse sur la tâche : de la 6ème à la 3ème, les élèves de SEGPA répondent nettement plus vite que des élèves de 6ème ordinaire. Une comparaison entre 6ème SEGPA et 6ème ordinaire montre que les élèves de SEGPA ne modulent presque pas leur temps de réponse en fonction de la difficulté des items : les élèves de 6ème ordinaire prennent nettement plus de temps pour répondre aux items difficiles, tandis que le temps de réponse des élèves de 6ème SEGPA est rapide et pratiquement constant au cours de la tâche. Enfin, on observe que cette rapidité de réponse sur les items difficiles n'est pas due à un désengagement face à la difficulté de la tâche : les élèves de SEGPA estiment en effet avoir très bien réussi la

tâche, au même degré que les élèves de classe ordinaire, en dépit du fait que leur performance est très nettement inférieure.

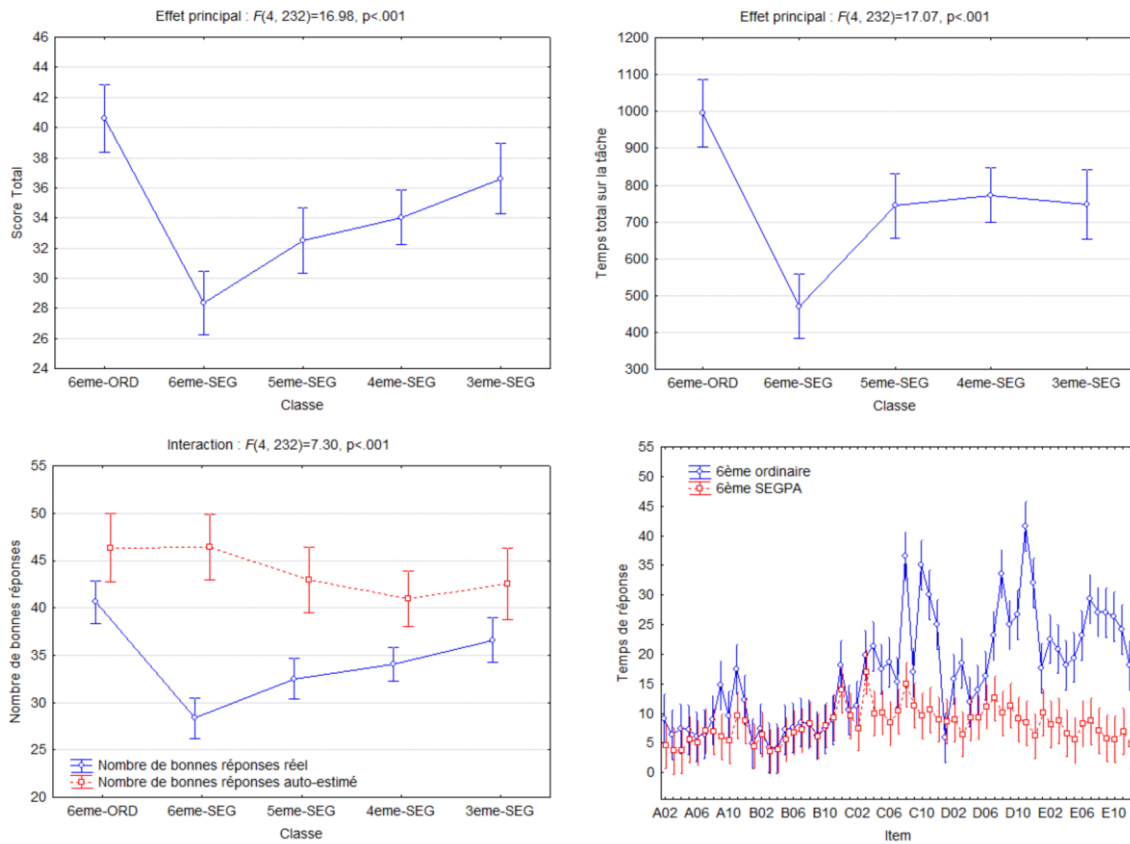


Figure 14. Score total, temps de réponse moyen total, nombre de bonnes réponses réel et auto-estimé, et temps de réponse moyen par item sur les matrices standard de Raven pour des élèves de 6ème ordinaire (ORD) et de 6ème-3ème SEGPA (SEG).

Ces résultats illustrent l'intérêt d'examiner les processus amenant à la réussite au niveau individuel, plutôt qu'au simple niveau agrégé de la performance totale. Examiner uniquement la variable inter-individuelle du score total amènerait à conclure que les élèves de SEGPA sont tout simplement moins intelligents, en moyenne, que les élèves de classe ordinaire. Mais une analyse plus fine des processus de résolution montre bien que cet écart de performance va de pair avec une différence plus profonde dans la façon de résoudre la tâche, les élèves de SEGPA étant loin de consacrer le temps nécessaire à la résolution des problèmes logiques. Nos résultats suggèrent que cette différence pourrait être due à un déficit dans l'aptitude métacognitive à évaluer la difficulté des problèmes, et donc le temps à consacrer pour les résoudre ; mais d'autres explications peuvent être envisagées, comme par exemple un besoin de cognition plus faible chez les élèves de SEGPA.



4. Le rôle des connaissances dans les tests d'intelligence

Les connaissances du participant ont une place importante dans les tests d'intelligence. La façon dont les connaissances contribuent à une tâche donnée n'est pas toujours évidente, mais les modèles systémiques de l'intelligence intègrent à peu près tous une contribution des connaissances acquises et/ou de la mémoire à long terme (voir Figure 10, page 124). Certains tests cherchent explicitement à évaluer les connaissances du participant ; c'est le cas des épreuves conçues pour mesurer l'intelligence cristallisée. D'autres tests cherchent au contraire à minimiser autant que possible le rôle des connaissances ; c'est le cas des épreuves dites *culture-free* ou *culture-fair*, dont la littérature a montré qu'elles ne sont jamais aussi indépendantes du bagage culturel qu'elles le souhaiteraient (e.g. Arvey, 1972 ; Kidd, 1962). Enfin, certains tests sont relativement agnostiques au sujet du rôle des connaissances, et intègrent des items sans s'apercevoir ou sans préciser qu'ils sont fortement dépendants du niveau de connaissances du participant ; certains tests de raisonnement quantitatif se placent dans cette catégorie.



On peut distinguer le rôle des connaissances déclaratives et des connaissances procédurales. Les connaissances déclaratives jouent directement un rôle dans les tests qui demandent explicitement au participant de faire appel à des connaissances acquises culturellement, souvent dans le cadre de la scolarisation : dans l'exemple des échelles de Wechsler, on peut citer par exemple les subtests de vocabulaire (définir un mot), de similitudes (indiquer le lien logique entre deux concepts), d'information (répondre à une question de culture générale), de compréhension (trouver une justification à une règle culturelle ou sociale), et d'arithmétique (résoudre un problème mathématique en utilisant des opérations connues qui font appel, par exemple, aux règles de proportionnalité). Cette dépendance aux connaissances acquises explique sans doute en grande partie le fait que la performance à ce type de tests décline peu avec l'âge (e.g. Baxendale, 2011, Grégoire, 1993), et le fait qu'elle corrèle fortement avec le niveau d'éducation et le niveau socio-économique (Dori & Chelune, 2004; Heaton et al., 2003; Holdnack et al., 2013). Cela explique sans doute également en partie les différences observées entre pays de cultures différentes (Lynn, 2015).

Les connaissances procédurales jouent un rôle dans tous les tests de raisonnement, quel que soit leur contenu, et même lorsqu'ils sont basés sur du matériel visuo-spatial abstrait comme les matrices de Raven. On peut distinguer deux grandes contributions des

connaissances procédurales. D'une part, les connaissances procédurales liées à la manipulation des items dans leur format particulier de présentation : un excellent exemple est le sens de lecture spontané, qui varie en fonction de la langue (de gauche à droite ou de droite à gauche, en arabe ou en hébreu par exemple), et du point de vue duquel les tests ne sont pas neutres – les matrices de Raven par exemple sont construites avec la logique d'une progression de gauche à droite, ce qui peut interférer avec la passation (Van de Vijver, 2016). Les tests de raisonnement quantitatif demandent une habitude dans la manipulation de chiffres, etc. Ce type de connaissances procédurales s'étend plus généralement à l'expertise dans la manipulation du contenu abstrait : certains participants ont moins l'habitude que d'autres de raisonner avec du matériel qui ne correspond pas directement à des objets concrets (Nunes Carraher, 1985 ; Flynn, 2007), ou peuvent même considérer le raisonnement abstrait comme moins pertinent que le raisonnement sur des propriétés concrètes (Cole et al., 1971 ; voir Greenfield, 1997) ce qui peut largement contribuer à une performance plus faible. D'autre part, les connaissances procédurales liées à la situation de test elle-même : les participants ont une expertise variable quant à la bonne façon de passer un test (connaissances liées par exemple à la façon de gérer son temps, au fait qu'il n'existe qu'une bonne réponse, etc ; Sternberg, 1999), mais aussi quant à la bonne façon de négocier une situation d'interaction dyadique avec un adulte qu'ils ne connaissent pas, de façon à résoudre un problème seul et sans l'appui d'autres membres de leur communauté (Greenfield, 1997).

La prise de conscience du rôle des connaissances déclaratives et procédurales dans les tests d'intelligence, dans une perspective constructiviste, amène à largement repenser les résultats ayant trait à des différences de score global, en particulier dans le contexte de différences inter-groupes. En effet, un score d'intelligence plus faible ou plus élevé ne s'interprète pas de la même façon si on considère qu'il est dû à une différence intrinsèque dans le niveau du participant sur une entité causale préexistante, d'origine génétique par exemple ; ou bien si l'on considère qu'il est dû à une différence de connaissances, acquises dans un environnement qui peut être plus ou moins favorable au développement de ces connaissances. Cette différence d'interprétation est particulièrement critique dans la mesure où elle a bien sûr de très importantes conséquences sociétales. L'application la plus directe de ce raisonnement réside évidemment dans les différences raciales d'intelligence : les Bushmen sud-africains ont un QI moyen de 55 qui les place en moyenne en situation de déficience intellectuelle (Lynn, 2015), mais ce résultat peut évidemment s'interpréter comme une différence dans les



connaissances déclaratives et/ou procédurales dont ils disposent pour traiter le contenu d'un test conçu par des occidentaux. J'ai décliné ce raisonnement à deux autres situations.

Je me suis d'abord intéressé à un phénomène qui a fait l'objet d'une importante couverture médiatique en France : le déclin de l'intelligence. Ce déclin a été décrit sous la forme d'un "effet Flynn négatif" (Dutton & Lynn, 2015), c'est-à-dire d'une inversion de l'augmentation de la performance moyenne d'intelligence d'une génération à l'autre appelée "effet Flynn" (Flynn, 1984 ; Rundquist, 1936). La conclusion selon laquelle l'intelligence décline en France est intégralement basée sur une unique étude de Dutton et Lynn (2015), qui n'ont pas réalisé de collecte de données mais ont seulement réinterprété des données présentées dans le manuel de la WAIS-IV. Dans le cadre de la validation de la WAIS-IV, un échantillon de 79 participants a passé à la fois la WAIS-III et la WAIS-IV de façon à pouvoir vérifier la corrélation entre les deux versions. Dutton et Lynn constatent que le score étalonné de cet échantillon est plus élevé sur la WAIS-IV (par rapport à l'échantillon normatif collecté en 2009) que sur la WAIS-III (par rapport à l'échantillon normatif collecté en 1999) ; puisque ces 79 participants sont comparativement meilleurs par rapport à l'échantillon de 2009 que par rapport à l'échantillon de 1999, Dutton et Lynn concluent que l'aptitude moyenne a diminué entre 1999 et 2009 - soit un déclin de l'intelligence moyenne. D'autres auteurs, Woodley of Menie et Dunkel (2015) supposent que ce déclin est "certainement d'origine biologique".

Ce soi-disant déclin de l'intelligence a été fortement publicisé en France, mais il y a plusieurs raisons d'être sceptique à son sujet. Sans parler de l'étude elle-même, la première raison qui me pousse au scepticisme est la forte orientation idéologique des auteurs : Edward Dutton est docteur en théologie et non en psychologie, régulièrement associé à des positions eugénistes et/ou d'extrême-droite, et éditeur en chef d'une revue suprémaciste blanche ; Richard Lynn présente un positionnement similaire et est particulièrement connu pour ses thématiques de recherche - "les femmes sont moins intelligentes que les hommes", "les autres races sont moins intelligentes que les blancs" et "les nouvelles générations sont moins intelligentes que les anciennes". Leur publication discute, comme mécanismes pouvant expliquer le plus probablement le déclin de l'intelligence, une trop forte immigration en France de populations non-blanches faiblement intelligentes, et la tendance des gens peu intelligents à faire un plus grand nombre d'enfants. La seconde raison d'être sceptique est que le consensus actuel propose plutôt que l'effet Flynn est ralenti ou interrompu dans les pays développés

(Pietschnig & Voracek, 2015 ; Wongupparaj et al., 2015 ; mais ce n'est pas le cas dans toutes les méta-analyses, cf. Trahan et al., 2014), mais pas qu'il s'inverse. Seules quelques études observent un effet Flynn négatif (Dutton et al., 2016), et ces quelques cas pourraient facilement s'expliquer par des oscillations aléatoires autour d'une trajectoire moyenne stable.

Quant à l'étude proprement dite, la faible taille d'échantillon (79 participants servant à généraliser une conclusion à "la population française"), le fait que les auteurs n'aient jamais eu accès aux données et se soient basés uniquement sur des statistiques descriptives présentées dans le manuel, et le fait qu'ils n'aient réalisé aucun test statistique pour vérifier la significativité des différences observées n'incitent pas à la confiance ; mais il est nettement plus problématique, dans une perspective constructiviste, de constater que la différence de score observée entre WAIS-III et WAIS-IV se limite à quelques subtests : essentiellement vocabulaire, information, compréhension et arithmétique. Autrement dit, uniquement des subtests qui font largement appel aux connaissances (déclaratives) du participant.



Les connaissances déclaratives étant par définition dépendantes du contexte culturel dans lequel évolue le participant, il n'est pas forcément absurde que la performance moyenne sur la version WAIS-III de ces subtests décline s'ils abordent des éléments de connaissance qui ne sont plus aussi représentés aujourd'hui qu'ils l'étaient en 1999 (e.g. Beaujean & Sheng, 2010 ; Brand, 1989 ; Kaufman, 2010). La version WAIS-III intègre par exemple des items sur les personnages célèbres du XXème siècle (retirés de la WAIS-IV), des calculs de proportionnalité exprimés en francs, une question sur le système douanier qui n'existe pratiquement plus en France, ou encore des items sur le rôle civique du parlement (remplacés par des items portant sur l'écologie dans la WAIS-IV). On conçoit facilement que des participants recrutés en 2009 ou 2019 réussissent moins bien ces items, sans que leur intelligence moyenne soit pour autant plus faible que celle des participants de 1999. On peut donc voir une éventuelle différence sous l'angle d'un biais d'item matérialisé par un fonctionnement différentiel des items, les mêmes items étant moins bien réussis par un échantillon récent à niveau d'aptitude équivalent.



On retrouve ici la nuance d'interprétation majeure qui résulte d'une vision réaliste de l'intelligence (il existe une entité causale qui cause la performance à tous les tests, qui sont donc relativement interchangeables) ou d'une vision constructiviste (il existe plusieurs déterminants de la performance, dont le poids relatif peut varier d'un test à l'autre). Une posture réaliste conduit à considérer que n'importe quelle différence de performance entre

deux groupes est une différence d'intelligence ; une posture constructiviste conduit à considérer que la différence de performance entre deux groupes est une différence dans l'un des mécanismes constituant la performance, qui peut refléter un vrai déficit de l'un des groupes (difficulté de raisonnement...) ou bien l'effet d'un mécanisme contextuel qui ne renseigne aucunement sur le niveau de compétence en tant que tel (menace du stéréotype...). Il s'agit en fait d'une autre façon de discuter le non-respect du postulat d'homogénéité locale : le poids relatif des connaissances déclaratives apparaît comme plus important dans certains subtests, et une différence observée entre deux groupes sur ces subtests uniquement invite à considérer que le score total au niveau inter-individuel n'est pas un résumé acceptable de l'aptitude - puisqu'il existe une inhomogénéité locale qui invite à considérer que seul un mécanisme impliqué plus fortement dans ces subtests est en cause.



Ce raisonnement nous a conduits à vérifier l'hypothèse d'un déclin de l'intelligence en France, en ré-analysant d'une part les données de comparaison entre WAIS-III et WAIS-IV prêtées par l'éditeur, et en collectant d'autre part un nouvel échantillon de réplication (Gonthier, Grégoire, & Besançon, en révision). Nous avons observé que l'échantillon collecté par l'éditeur était peu adapté à la recherche d'un effet Flynn négatif (auquel il ne se destinait pas) généralisable à l'ensemble de la population française, puisqu'il était composé de deux groupes d'âge discrets (30-34 ans et 55-63 ans), et d'une proportion anormale de participants en situation de déficience intellectuelle qui présentaient un score plancher sur certains subtests.

Nous avons surtout confirmé que la différence observée entre WAIS-III et WAIS-IV portait exclusivement sur les subtests saturant l'intelligence cristallisée, et que cette différence s'expliquait largement par un fonctionnement différentiel des items : environ un quart des items étaient significativement plus difficiles pour l'échantillon 2019 que pour l'échantillon 1999, à niveau d'aptitude équivalent (l'aptitude étant estimée à partir des autres subtests disponibles, tels que les matrices de Raven). Cette conclusion remet en cause l'idée d'un déclin général en France de l'intelligence, et en particulier du raisonnement logique ; il soutient au contraire l'idée d'un déclin spécifique des connaissances déclaratives, telles que mises en jeu par la WAIS-III. Quant à l'argument selon lequel le déclin observé serait d'origine biologique (Woodley of Menie & Dunkel, 2015), il s'effondre largement quand on remarque que la

quantité de déclin observée pour un subtest corrèle à .91 avec la charge culturelle de ce subtest et à .86 avec la taille de l'effet de la catégorie socio-professionnelle sur ce subtest.

J'ai développé un autre projet de recherche portant sur la spécificité du rôle des connaissances déclaratives dans les tests d'intelligence (Tourreix, Besançon, & Gonthier, en préparation). Les arguments précédents concernant le caractère unitaire ou non de la performance d'intelligence invitent à questionner le phénomène de haut potentiel intellectuel (HPI). Le HPI se définit de façon usuelle par un QI total supérieur à 130 ; mais quels mécanismes sont à l'origine de cette performance globalement élevée, et dans quelle mesure cette performance élevée est-elle homogène ? Ces questions ne se poseraient pas si l'intelligence constituait une entité réaliste unitaire ; mais si plusieurs mécanismes différents contribuent à la performance et si leurs poids respectifs varient d'une tâche à l'autre, on peut s'attendre à ce que des participants d'aptitude élevée montrent des patterns de forces et de faiblesses particuliers, donc une certaine hétérogénéité - puisqu'il n'existe pas, par définition, de mécanisme central fortement déficitaire qui produirait une contrainte forte sur la performance globale (Kovacs & Conway, 2016).



Plusieurs études se sont penchées sur les profils de scores des personnes HPI, et ont relevé une hétérogénéité plus importante des performances que chez des participants d'aptitude plus faible (e.g. Fishkin et al., 1996 ; Hollinger, 1986 ; Silver & Clampit, 1990). Cette hétérogénéité est souvent décrite comme une dyssynchronie développementale, à caractère pathologique (e.g. Guénolé et al., 2015), mais s'explique en réalité très bien dans le contexte de déterminants multiples de la performance qui ne sont pas uniformément efficaces et agissent de façon différente sur des tâches différentes (e.g. Labouret & Grégoire, 2018). Ce phénomène est résumé par la loi des rendements décroissants de Spearman, c'est-à-dire le fait que la corrélation entre les tâches tend à diminuer lorsque le niveau d'aptitude augmente (Blum & Holling, 2017 ; Detterman & Daniel, 1989 ; Spearman, 1927).

Le point qui m'intéresse ici est que cette hétérogénéité n'est pas forcément aléatoire, et qu'elle n'est pas neutre du point de vue des mécanismes en jeu. On retrouve souvent un profil de score particulier dans lequel les personnes HPI présentent une performance nettement supérieure à l'indice de compréhension verbale des échelles de Wechsler (vocabulaire, information, compréhension, similitudes ; voir page 143), qui reflète largement le rôle des connaissances déclaratives (pour un exemple, voir Labouret & Grégoire, 2016). Dans la mesure

où ce type de connaissances est fortement dépendant du milieu culturel et socio-économique dans lequel évolue la personne, on peut s'attendre à ce que ce type de profil indique non pas un très haut niveau d'aptitude intellectuelle (incluant, par exemple, de fortes compétences de raisonnement fluide), mais plutôt l'appartenance à une classe sociale élevée. Il pourrait donc exister de "faux HPI", qui atteignent le seuil de QI suffisant pour être classés comme tels du seul fait de connaissances déclaratives très élevées qui conduisent à surestimer leur performance globale. (Cela n'exclut pas, bien sûr, que les personnes HPI présentent légitimement d'excellentes connaissances cristallisées, par investissement d'excellentes aptitudes fluides ; mais un profil de type *Indice de Compréhension Verbale = 130, Indice de Raisonnement Perceptif = 100* nous mettrait plutôt en présence d'une surestimation éventuellement liée à la classe sociale.)



La littérature fournit peu de renseignements sur ce point : les quelques études qui ont examiné les patterns de scores de participants à haut potentiel (e.g. Fishkin et al., 1996 ; Hollinger, 1986 ; Silver & Clampit, 1990) ont généralement considéré le HPI comme une classe diagnostique homogène et ont donc uniquement examiné les scores moyens sur différents types de tâches, sans chercher à repérer de sous-profils. Il existe une littérature très limitée concernant la distinction entre les profils équilibrés et les profils de type "QI verbal > QI performance" ou inversement : Wilkinson (1993) par exemple trouve que les participants dont le "QI verbal" (qui reflète la performance aux subtests intégrant un rôle important des connaissances déclaratives) est nettement supérieur à leur "QI performance" tendent à avoir un niveau académique significativement plus élevé - ce qui va bien dans le sens d'un profil qualitativement différent car reposant sur un mécanisme causal spécifique aux connaissances déclaratives.

Je ne crois pas qu'il existe de données publiées se rapportant à une analyse de profils formelle dans un échantillon de participants à haut potentiel. Nous avons donc examiné cette question (Tourex, Besançon, & Gonthier, en préparation) en nous appuyant sur un échantillon d'environ 300 protocoles WISC et WAIS de personnes diagnostiquées comme "à haut potentiel", et fournis par des psychologues cliniciens exerçant en libéral. Une analyse de profils, dont les résultats sont résumés dans la Figure 15, a permis de montrer l'existence de deux classes distinctes de personnes à haut potentiel. La première classe est caractérisée par une performance nettement plus importante sur l'indice de compréhension verbale que sur tous les

autres indices, ce qui correspond à l'idée d'une surestimation de la performance intellectuelle globale du fait de connaissances déclaratives significativement supérieures à la norme. La seconde classe est caractérisée par une performance relativement homogène à travers l'ensemble des indices (la performance apparemment plus faible pour cette classe sur l'indice de compréhension verbale dans la Figure 15 est un artefact, dû à l'usage de scores standardisés au sein de cet échantillon).

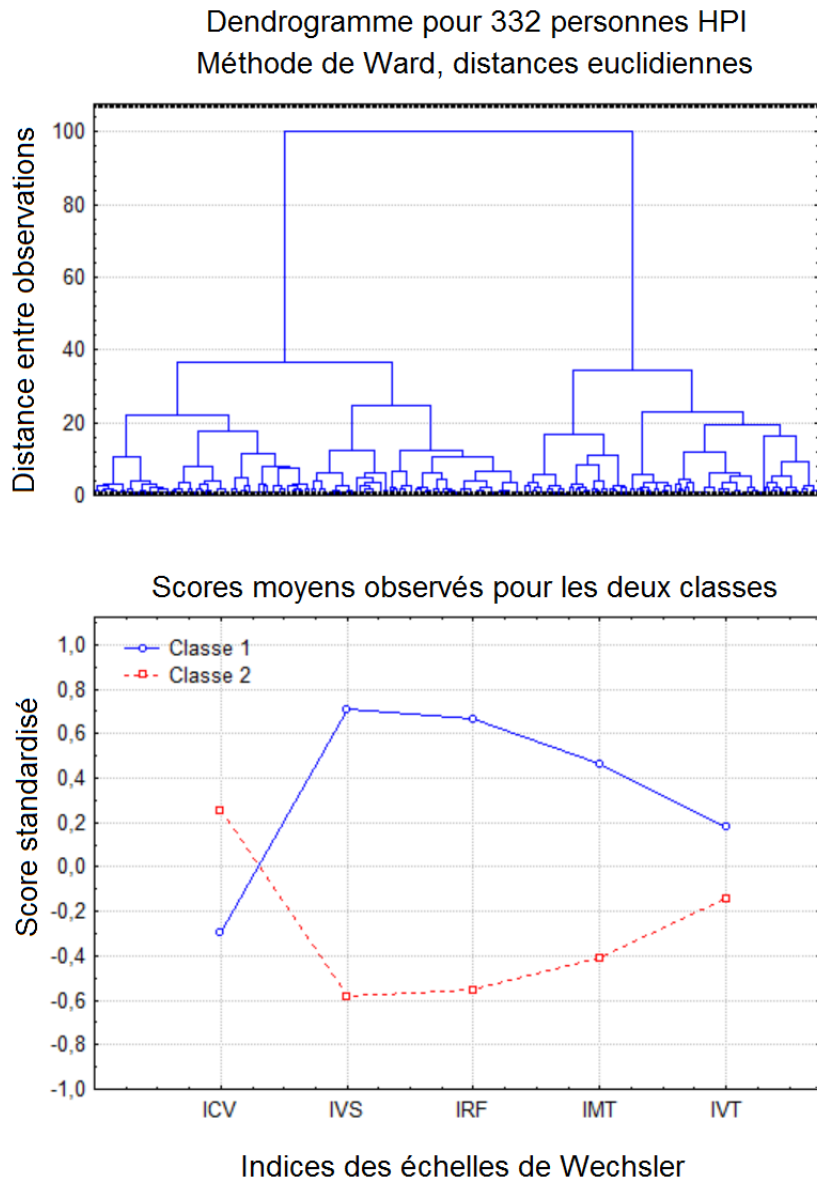


Figure 15. Analyse de profils portant sur 332 protocoles WISC ou WAIS de personnes à haut potentiel. En haut : dendrogramme extrait dans le cadre d'une analyse de classification hiérarchique ; en bas : scores moyens aux cinq indices des échelles de Wechsler pour les participants affectés à l'une de ces deux classes dans le cadre d'une analyse en k -moyennes.

Ces résultats suggèrent qu'il existe bien deux profils de personnes HPI : un profil caractérisé spécifiquement par une performance élevée exclusivement sur les subtests qui dépendent des connaissances déclaratives, et un profil caractérisé par une performance élevée de façon homogène, notamment sur les autres types de subtests. Cette conclusion a bien sûr des implications sociétales : le premier profil est susceptible d'être surreprésenté au sein des classes sociales les plus élevées, et les parents d'un niveau socio-économique plus élevé ont de plus tendance à être surreprésentés au sein des échantillons qui consultent pour faire poser un diagnostic de haut potentiel à leurs enfants.

Ce constat soulève également des questions concernant la pratique récente consistant à diagnostiquer un haut potentiel dès lors que l'un des indices d'une échelle de Wechsler est supérieur à 130 (on retrouve par exemple cette politique chez les professionnels du Centre National d'Aide aux enfants à Haut Potentiel). L'argument avancé est de mieux reconnaître les spécificités et les forces et faiblesses de chaque enfant, mais cette pratique risque fort de conduire à un sur-diagnostic de HPI chez les enfants de classes sociales supérieures qui tendent à présenter un ICV plus élevé. L'absence d'une entité réaliste "intelligence" et l'existence d'inhomogénéités locales au sein des profils de score conduisent ainsi à s'interroger sur la pertinence d'une catégorie diagnostique unique pour le haut potentiel : non seulement parce qu'une décomposition en "haut potentiel de raisonnement" et "haut potentiel académique" pourrait avoir plus de sens, mais aussi plus généralement parce que l'absence d'une entité causale unique doit amener logiquement à préciser quels sont les mécanismes en jeu dans la détermination de la haute performance de chaque individu.

5. Perspectives de recherche

5.1. Sur le rôle des stratégies dans les tests d'intelligence

Perspective 3.1.A : mes résultats concernant la médiation du lien entre capacité en mémoire de travail et intelligence par les stratégies (Gonthier & Thomassin, 2015) me paraissent inviter une réplication et/ou une extension. D'une part, la non-réplication de ces résultats par Jastrzębski et collaborateurs (2018) est problématique. D'autre part, la méthode que j'ai utilisée pour induire le constructive matching, en étayant le processus de résolution à travers une décomposition des étapes de traitement à effectuer, ne paraît pas très satisfaisante car elle est susceptible d'affecter d'autres facteurs que la stratégie employée. On peut penser notamment au rôle du contrôle cognitif (décomposition du processus de résolution en une hiérarchie de buts et de sous-but à traiter séquentiellement ; Carpenter et al., 1990).

En ce qui concerne l'évaluation des stratégies, ces travaux gagneraient à être étendus à d'autres types de tâches, les matrices de Raven étant la tâche d'intelligence la plus employée en recherche mais certainement pas la seule. Les cubes de Kohs (Rozencwajg, 1991 ; Rozencwajg & Corroyer, 2001) ou les tâches de syllogismes par exemple se prêteraient bien au test de la même hypothèse. En ce qui concerne la manipulation expérimentale des stratégies, d'autres solutions existent pour induire le constructive matching, dont les plus directes consistent à contraindre le participant à regarder la matrice un certain temps avant de voir les réponses possibles (Rivollier, Quinton, Gonthier, & Smeding, 2020), ou tout simplement à ne pas proposer de réponses possibles et à demander au participant de dessiner directement la bonne réponse (Becker, Schmitz, et al., 2016 ; Mitchum & Kelley, 2010). Ces deux approches contraignent ou empêchent l'utilisation de l'élimination de réponses et devraient donc augmenter la corrélation avec la capacité en mémoire de travail (si tant est que diminuer l'élimination de réponses augmente mécaniquement le constructive matching ; mes résultats sur l'effet de la pression temporelle, voir page 140, suggèrent que les choses ne sont pas aussi simples et que l'on risque plutôt d'augmenter le guessing).

Perspective 3.1.B : la mise en évidence du rôle des stratégies comme mécanisme causal de la performance dans les matrices de Raven, mais aussi comme mécanisme faisant l'objet de différences inter-individuelles, conduit à questionner dans quelle mesure la variabilité stratégique contribue à la variabilité du score total. Cette question peut se décliner à n'importe

quelle comparaison entre deux échantillons mais me paraît particulièrement intéressante dans deux contextes.

Le premier est celui du HPI : quelques études ont montré que les enfants HPI présentaient des comportements stratégiques plus efficaces que des enfants contrôles (Coyle et al., 1998 ; Gaultney et al., 1996), mais cette question semble étonnamment ne pas avoir été examinée dans le cadre des tests d'intelligence. Montrer que les personnes HPI tendent à utiliser des stratégies plus efficaces à travers différents types de tests pourrait fournir un élément contribuant à expliquer leur performance intellectuelle supérieure.

Le second est celui de la trajectoire développementale de l'intelligence fluide. On sait que la performance sur les tests d'intelligence fluide comme les matrices de Raven augmente considérablement jusqu'à l'âge adulte, puis diminue au cours du vieillissement normal (e.g. Baxendale, 2011 ; Raven, 2000). Une interprétation simpliste de ces résultats consisterait à dire que les humains deviennent de plus en plus intelligents en grandissant puis de plus en plus bêtes en vieillissant, ce qui confine à l'absurde : il serait nettement plus intéressant de préciser les mécanismes en jeu dans cette trajectoire développementale et à l'origine de la variabilité des scores. Ces mécanismes sont certainement multiples (vitesse de traitement, mémoire de travail, connaissances procédurales... évoluent avec l'âge et peuvent contribuer à cette variation), et la variabilité stratégique peut en faire partie (e.g. Rozenchwajg & Corroyer, 2001 ; Rozenchwajg et al., 2005). Il serait intéressant de vérifier dans quelle mesure la trajectoire vie entière des performances de raisonnement s'accompagne de changements dans les stratégies utilisées.



Dans le cas des matrices de Raven, il est possible de tester cette hypothèse même chez de jeunes enfants pour lesquels l'auto-report de la stratégie utilisée est trop difficile d'un point de vue métacognitif (j'en ai fait l'expérience) : le protocole de décomposition des matrices de Raven sur lequel nous avons travaillé s'y prête bien (Rivollier, Quinton, Gonthier, & Smeding, 2020 ; voir pages 135-136). J'ai mené une première étude en ce sens, dont les résultats préliminaires sont extrêmement prometteurs ; ils sont représentés dans la Figure 16. Il reste à étendre ce protocole à une population d'adultes âgés pour vérifier que la covariation de la performance et du constructive matching se poursuit dans le même sens dans le cas du vieillissement normal ; à rechercher les processus pouvant contribuer à expliquer cette évolution stratégique (la capacité en mémoire de travail est un candidat possible, tout comme



le besoin de cognition ; mais l'aptitude métacognitive à évaluer la difficulté des problèmes en est un autre, cf. Perret & Dauvier, 2020). Et enfin, il faudra chercher à répondre à la question du rôle causal des stratégies dans cette évolution : l'augmentation progressive de l'usage du constructive matching est-elle la cause de l'augmentation développementale de la performance de raisonnement, ou en est-ce un symptôme (les enfants les plus aptes à reconstruire la structure logique de la matrice utilisent plus cette stratégie de reconstruction) ?

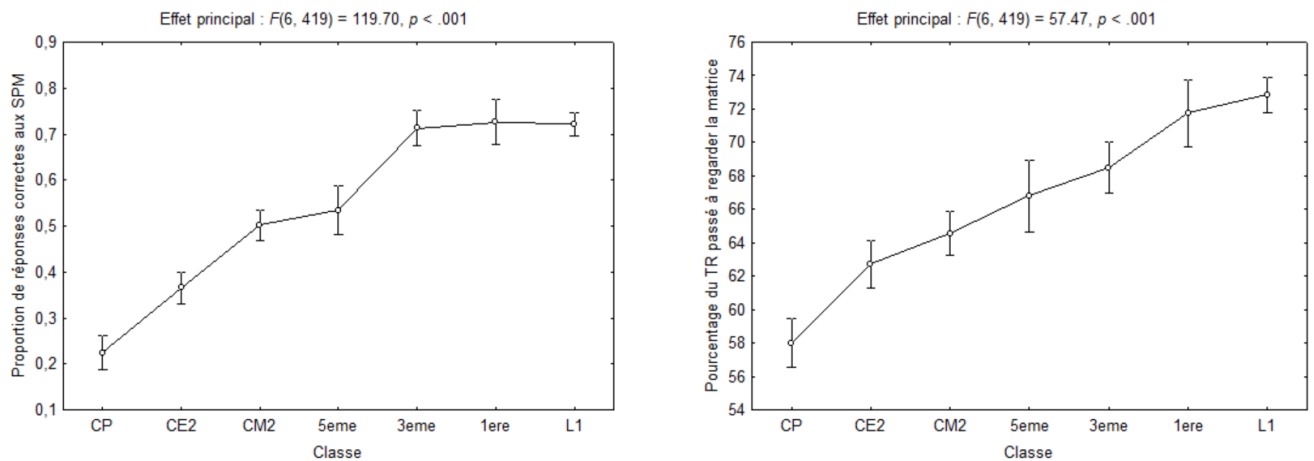


Figure 16. Trajectoire développementale de la performance aux matrices standard de Raven (à gauche) et de l'usage du constructive matching, représenté par le pourcentage de temps passé à regarder la matrice avant de regarder les réponses possibles (à droite). Le constructive matching médie partiellement l'effet de l'âge sur la performance, $z = 9.10$, $p < .001$.

Perspective 3.1.C : le constat qui m'a frappé dans le cadre de mes travaux sur les stratégies est la faiblesse de nos connaissances quant aux stratégies utilisées dans le cadre d'épreuves cliniques. La variabilité stratégique semble fortement contribuer à la variabilité de la performance : pourquoi ne prendrait-on pas en compte cette variable dans l'évaluation de l'intelligence ? C'est peut-être que l'on ne sait pas toujours quelles stratégies peuvent être utilisées sur les différents subtests d'une échelle de Wechsler. Ce travail a été fait pour les tests de matrices, de cubes, d'arithmétique ou de mémoire de travail ; il ne l'a pas du tout été pour les tests de vocabulaire, de similitudes, de balances, de puzzles visuels. Un projet de recherche portant sur l'identification des stratégies possibles dans ce type de tests et sur leur contribution à la performance me paraît donc absolument nécessaire, de façon similaire à mon travail de recensement des stratégies possibles en mémoire visuo-spatiale (voir pages 93 et 110), avec des applications cliniques possibles.



5.2. Sur les autres déterminants de la performance

Perspective 3.2.A : mes travaux sur le développement d'une tâche de mémoire de travail adaptative (page 81) trouveraient une application logique dans la conception de tests d'intelligence. Rien n'empêche en effet d'envisager une tâche d'intelligence adaptative, dans laquelle on commencerait par un item de difficulté intermédiaire pour ensuite moduler la difficulté des items présentés en fonction de la performance du participant. Ce type de tâche présenterait les mêmes avantages que dans le domaine de la mémoire de travail : possibilité de raccourcir les passations de tests, risque moindre d'interrompre une passation trop tôt lorsque le critère d'arrêt d'une procédure ascendante est atteint par désengagement ou inattention du participant, mais aussi possibilité de proposer un plus grand nombre d'essais réellement adaptés au niveau du participant de façon à améliorer la précision de la mesure.

La procédure adaptative serait simple à mettre en œuvre pour les tests portant plutôt sur l'intelligence cristallisée ; elle est d'ailleurs très souvent utilisée dans les tests d'aptitude académique dans le monde anglo-saxon, dont la structure est très similaire. Elle soulève en revanche quelques problèmes intéressants dans le cas des tests de raisonnement. Dans le cas d'un test comme le D2000 ou les matrices de Raven, la procédure ascendante peut être vue comme un moyen d'appréhender progressivement la structure de la tâche et les règles logiques qu'elle implique : les mêmes règles sont réutilisées à des niveaux différents de complexité (voir Carpenter et al., 1990). On peut ainsi douter de la capacité d'un participant à inférer une règle logique si elle lui était présentée pour la première fois sur un item au plus haut niveau de difficulté - alors que la complexité perceptive et le nombre d'éléments logiques de la matrice sont bien plus importants que dans le cas des items plus simples. En court-circuitant les items les plus faciles pour les participants d'aptitude élevée, l'approche adaptative pourrait donc interférer avec la bonne résolution de la tâche. Mais cette limite n'est pas rédhibitoire et on peut imaginer des solutions de contournement, consistant par exemple à présenter systématiquement une série d'items simples servant à introduire les différentes règles logiques avant d'entamer la phase adaptative proprement dite, ou même indiquer directement au participant en début de tâche quelles règles logiques il est susceptible de rencontrer, ce qui ne semble pas modifier radicalement la nature de l'épreuve (Loesche et al., 2015).



Perspective 3.2.B : les connaissances procédurales sont peut-être le déterminant de la performance de raisonnement qui a été le moins étudié. Elles sont pourtant à mon avis l'un des

plus intéressants, notamment dans le contexte des différences entre groupes ethniques (Lynn, 2015). On sait que le sens de lecture a une influence sur les tests visuo-spatiaux en général, par exemple dans le cadre des études portant sur la ligne numérique (e.g. Zohar-Shai et al., 2017) ou la perception de l'espace (Fagard & Dahmen, 2003) ; il devrait également poser problème sur des tests de raisonnement comme les matrices de Raven. Il en existe des exemples anecdotiques dans la littérature (Sternberg & Rifkin, 1979 ; van de Vijver, 2016) mais je crois qu'aucune étude systématique n'a jamais été publiée à ce sujet. Il serait intéressant de comparer une condition standard et une présentation inversée en miroir des matrices de Raven, au sein d'échantillons français et de langue arabe ; certains pays comme le Maroc se prêtent bien à ce type de recherche car ils mélangent des populations parlant français, arabe ou les deux, à des degrés différents dans la vie quotidienne.



5.3. Vers une architecture causale du raisonnement (CAR)

Les travaux menés en psychologie cognitive et en psychologie différentielle depuis un peu plus d'un siècle ont permis de considérablement développer nos connaissances sur la performance humaine dans les tests d'intelligence. Je crois malheureusement que le grand échec de nos disciplines est le manque de transcription directe de ces travaux dans la représentation quotidienne des psychologues.

La position réaliste sur l'intelligence n'a jamais été une proposition sérieuse de la part de ses théoriciens, qui se sont tous ou presque accordés à voir l'intelligence comme le résultat d'un ensemble de processus en interaction. Des modèles systémiques de l'intelligence en ont découlé, mais ils sont restés très confidentiels et sont apparus essentiellement dans des chapitres d'ouvrages, même lorsqu'ils étaient portés par les auteurs les plus reconnus (Kyllonen, 2002 ; Schneider & McGrew, 2012). Au final, la vision constructiviste de l'intelligence, pourtant assez consensuelle pour les chercheurs, apparaît comme étant plutôt confinée aux travaux de quelques spécialistes.

Mon sentiment est que de cette pathologie découlent des symptômes à trois niveaux :

- 1) l'intelligence est généralement enseignée aux jeunes psychologues sous l'angle simpliste du facteur général, puis sous l'angle des modèles factoriels, en omettant le lien possible avec l'architecture cognitive dans son ensemble et avec la complexité des autres déterminants de la performance qui ne génèrent pas forcément de facteurs statistiques ;
- 2) les tests restent conçus pour tirer une interprétation globale au niveau d'un score total, et les psychologues cliniciens sont avant tout entraînés à tirer ce type d'interprétation ; le critère d'homogénéité des indices pour pouvoir calculer un score total, ou encore l'encouragement par le manuel de la WAIS à considérer l'importance de déterminants socio-motivationnels de la performance, ne suffit pas à y remédier et nous voyons régulièrement des exemples de dérives diagnostiques, une dyspraxie étant confondue avec une déficience intellectuelle ;
- 3) les chercheurs non-spécialistes ou non intéressés par les travaux portant sur l'origine causale des scores d'intelligence l'utilisent comme une entité réaliste pour tirer des interprétations tout aussi simplistes sur les différences inter-individuelles, les travaux de Lynn (2015) sur les différences raciales constituant un excellent exemple de cette dérive. On pourrait ajouter un quatrième problème qui aurait trait à la représentation que le grand public se fait de l'intelligence, mais je crois ce combat largement perdu d'avance.

Je pense que cette faible popularité est due, en partie, à la faible visibilité des travaux portant sur la multi-détermination des scores d'intelligence. La question fondamentale de la posture épistémologique à adopter est généralement présentée dans des travaux très complexes et dont la discussion porte en bonne partie sur la modélisation statistique (oui, Molenaar, 2004, c'est toi que je regarde). La multiplicité des déterminants en jeu est éparpillée dans des articles s'intéressant chacun à une variable et sans lien apparent entre eux. L'approche systémique, à mon avis la plus fructueuse comme cadre d'analyse transversale, a reçu très peu d'attention depuis sa formulation initiale il y a une quarantaine d'années (Detterman, 1982, 1987), et les modèles qui en découlent sont peu accessibles, généralement relégués à des chapitres d'ouvrages. Tout cela contribue à dissimuler l'existence d'une alternative à la vision réaliste du score global aux non-spécialistes, et même aux spécialistes qui ne la cherchent pas explicitement. Je crois d'ailleurs assez parlant que l'un des principaux reproches adressés à la *process overlap theory* de Kovacs et Conway (2016) soit le fait qu'ils ne fassent que redécouvrir une position constructiviste déjà bien formulée par ailleurs (Gottfredson, 2016) ; on n'accusera pourtant pas Andrew Conway d'en être à ses premiers travaux dans le domaine de l'intelligence.

À ce stade, je crois que le plus grand service que l'on puisse rendre à ce champ de littérature serait la synthèse d'un modèle qui représente, de façon relativement exhaustive, l'ensemble du système de mécanismes fonctionnels - aussi bien quantitatifs que qualitatifs - dont le déploiement est requis pour réussir les tests d'intelligence ; accompagné de la formulation explicite des conséquences de cette vision constructiviste, notamment la multi-détermination, la multi-réalisabilité, et la possibilité de contributeurs multiples aux corrélations avec d'autres variables. Un tel système de référence, dans une revue à forte visibilité, pourrait donner un point d'ancrage à l'enseignement des mécanismes de l'intelligence, à l'interprétation clinique d'un faible score d'intelligence, et à la recherche des mécanismes potentiellement impliqués dans les différences entre groupes, entre individus, ou entre classes d'âges.

Je doute que l'on découvre encore des mécanismes radicalement nouveaux du fonctionnement de l'intelligence : certains ont été plus étudiés que d'autres (le raisonnement inductif ou déductif en tant que tel, c'est-à-dire en tant que processus itératif consistant à formuler des hypothèses et à vérifier leur application, est peut-être paradoxalement le moins bien connu de ces mécanismes), et on peut discuter des distinctions et chevauchements à



établir entre ces mécanismes, mais la liste que l'on peut en faire est sans doute relativement complète, et les systèmes cognitifs proposés dans la Figure 10 (page 124) constituent sans doute une approximation valable. Ces systèmes manquent toutefois profondément de raffinement dans la perspective d'une utilisation clinique. J'aimerais évoquer ici une dernière approche : celle de Snow (1992), qui porte sur l'aptitude en général mais se décline parfaitement au domaine de l'intelligence, et répond largement à ces limites.

La synthèse de Snow (1992) sur les aptitudes présente quelques originalités. D'une part, elle insiste sur la multiplicité des déterminants en jeu, en incluant plusieurs déterminants qui ne sont presque jamais discutés dans le champ de l'intelligence. Snow donne les exemples présentés dans la Figure 17, qui s'appliquent à l'apprentissage mais se déclinent bien également aux tests d'intelligence : on y retrouve des déterminants émotionnels et motivationnels absents des systèmes cognitifs de la Figure 10, et généralement sous-estimés par les chercheurs cognitivistes (Raven, 2008), mais qui doivent jouer un rôle dans les scores d'intelligence et particulièrement en contexte clinique : la motivation (Duckworth et al., 2011 ; que l'on pourrait décliner au besoin de cognition : Hill et al., 2013, 2016), le besoin d'accomplissement, ou encore l'anxiété liée au test (Oostdam & Meijer, 2003 ; Meijer & Oostdam, 2007). Les comportements stratégiques, bien connus dans la littérature cognitive mais absents des modèles systémiques, font d'ailleurs partie de cette liste (Snow est d'ailleurs le premier auteur à identifier les deux stratégies de constructive matching et d'élimination de réponses).

<i>Aptitude Category</i>	<i>Aptitude Constructs</i>		
	<i>Stable?</i>	<i>Malleable?</i>	<i>Ephemeral?</i>
Conceptual structures	Crystallized intelligence	Prior knowledge of physics	Believes concept X Explains concept X
Procedural skills	Fluid intelligence	Quantitative ability	Computes automatically
		Selective attention	Infers relations Detects cues
Learning strategies	Visualization ability	Imagery	Graphs relations
		Representation of goals	Tests hypotheses
Self-regulatory functions	Deep versus surface processing	Heuristics, and mnemonics	Checks progress
		Adaptation of strategy	Works alone
Motivational orientations	Independence–flexibility	Action orientation	Protects intention
	Conformity–responsibility	Content motivation	Works extra
	Extraversion–ascendancy	Investigative interest	Explores resources
	Need for achievement	Mastery orientation	Seeks explanation
	Test anxiety	Self-orientation	Worries about failure

Figure 17. Déterminants multiples de la performance dans le domaine de l'apprentissage. Tiré de Snow, R. (1992), *Aptitude theory: Yesterday, today, and tomorrow*. *Educational Psychology*, 27(1), 5-32.

Par ailleurs, et même si ce n'est pas apparent dans la Figure 17, Snow insiste également sur le fait que ces déterminants doivent être vus sous la forme d'un système dont les composants interagissent - *human learners are not just lists of variables*. Cette idée rejoint les approches systémiques évoquées précédemment, à la fois sous l'angle d'influences réciproques possibles (van der Maas et al., 2006) et de dépendences fonctionnelles sous la forme de goulots d'étranglement (Detterman, 1987) ou de substitutions ; mais elle les complète également puisqu'elle invite à considérer de possibles interactions entre mécanismes cognitifs, motivationnels, émotionnels, etc. Plus qu'une théorie des déterminants potentiels de la performance, ce qui manque largement à la littérature contemporaine est une théorie des inter-relations entre ces déterminants : *we need theories of aptitude complexes, not just of single constructs* (Snow, 1992).

Enfin, Snow insiste sur la définition de l'aptitude (je dirais : de la performance) comme la réalisation de l'ensemble de ces déterminants *dans une situation donnée* : la performance d'un individu est indissociable de la situation dans laquelle cette performance s'exerce, les variables individuelles interagissant avec l'environnement dans lequel le participant est placé. Un participant n'est donc pas apte ou inapte, mais apte ou inapte dans une situation donnée ; c'est une conséquence logique de la multi-détermination lorsque les déterminants sont indépendants et que leur poids peut varier d'une situation à l'autre (page 122), ce qui remet bien sûr directement en cause le postulat d'homologie (page 9). Un participant malvoyant chutera dans un test visuel mais, évidemment, pas nécessairement dans un test verbal.

Cette prise en compte de la situation ne se comprend pas uniquement sous l'angle d'interactions entre des variables individuelles et des variables situationnelles stables, liées à la nature de la tâche : certains des déterminants de la performance évoqués dans la Figure 16 sont purement éphémères et liés à des caractéristiques de la situation qui ne sont pas forcément utiles pour prédire la performance future du participant (le fait de travailler seul, de travailler dur, de s'inquiéter au sujet des conséquences d'un échec). Snow insiste également sur le caractère dynamique de ces interactions entre participant et environnement : les différents déterminants et leur organisation dans une tâche donnée peuvent varier d'un participant à l'autre, mais aussi au cours du temps. Le système n'a donc pas de définition unique ; dans une certaine mesure, le système de déterminants à l'origine de la performance ne pourrait être

parfaitement décrit que de façon idiographique - au niveau d'une tâche donnée pour un individu donné à un moment donné.

Ce degré de granularité est peut-être peu utile pour une recherche en psychologie avant tout préoccupée par la définition de lois générales (ce qui inclut les lois générales de la variabilité ; l'approche purement idiographique n'est de façon générale pas franchement compatible avec la généralisation requise par la recherche scientifique), mais il apparaît indispensable pour une utilisation clinique des tests d'intelligence, qui vise avant tout à comprendre le participant dans sa spécificité : un déterminant éphémère de la performance comme l'anxiété liée au test peut agir comme un goulot d'étranglement et jouer un rôle crucial dans la chute de performance d'un patient donné, qui ne doit pas pour autant faire conclure à une déficience intellectuelle.

Il reste donc à formuler explicitement la proposition d'un système de processus quantitatifs (cognitifs et socio-motivationnels) et qualitatifs (comportements stratégiques), et de la nature possible de leurs interactions, tels qu'ils s'inscrivent dans une situation particulière ; ce système étant considéré comme généralement applicable, sous un angle nomothétique, mais avec la possibilité que le poids relatif des différents mécanismes varie d'un participant à l'autre et d'une situation à l'autre. Cette description nomothétique est une condition du projet idiographique que constitue la psychologie clinique ; et rendre cette synthèse accessible sera la condition de son application pratique à l'enseignement, à la clinique et à la recherche.

La Figure 18 représente une ébauche de ce à quoi pourrait ressembler un tel système de mécanismes en jeu dans les tests d'intelligence. J'utilise une représentation similaire dans le cadre de mes enseignements de licence portant sur l'intelligence. Le titre, Causal Architecture of Reasoning (CAR), est un jeu de mots et un moyen mnémotechnique qui s'appuie sur une analogie souvent utilisée pour présenter l'approche constructiviste, celle d'un moteur de voiture (e.g. Borsboom, 2009 ; voir aussi Spearman, 1927) : ce système peut être vu comme un moteur, dont la performance dépend de l'ensemble du système et non d'un composant en particulier ; dont les composants sont fortement intégrés (Detterman, 1987) et pour lequel une quelconque atteinte peut provoquer une chute catastrophique de la performance ; et dont la performance dépend du contexte dans lequel on le place.

Comme tous les modèles, ce modèle est faux, évidemment. Plusieurs problèmes sont immédiatement apparents : certains concepts ne sont pas clairement définis ou se chevauchent avec les concepts voisins ; on pourrait poursuivre *ad nauseam* dans le réductionnisme et considérer que chacun des mécanismes représentés comme un déterminant de la performance est lui-même une propriété émergente d'un système de mécanismes en interaction. Les relations entre concepts sont souvent très hypothétiques (et pour cause, puisque l'absence de travaux conceptualisant explicitement les mécanismes de l'intelligence comme un système a mené à une quasi-absence d'études portant sur les interactions entre eux). Toutes les relations plausibles ne sont pas représentées ; la variabilité situationnelle est difficile à faire apparaître (peut-être en jouant sur la couleur des boîtes pour montrer de quelle façon leur poids relatif peut changer d'une tâche à l'autre ou au cours du développement) ; les stratégies sont représentées sous la forme d'un déterminant comme les autres mais un changement qualitatif de mode de traitement pourrait affecter la structure du reste du système. Cette proposition de modèle serait compliquée à tester ou falsifier dans son ensemble ; ce n'est d'ailleurs pas la bonne solution, à mon avis, puisque la seule option possible pour cela serait une modélisation statistique à grande échelle, qui ne peut pas être adaptée pour des mécanismes comme la perception qui ne sont pas nécessairement générateurs de variabilité.



Bref, les problèmes ne manquent pas, mais la valorisation d'une proposition de ce type serait à mon avis le meilleur moyen de diffuser l'état de nos connaissances concernant la complexité du système à l'origine de la performance d'intelligence, et d'en faire profiter l'enseignement, la clinique, et la recherche des non-spécialistes. À moyen terme, je pense que le moyen le plus fructueux de progresser dans cette direction serait la création d'une grille d'aide à l'interprétation des tests d'intelligence dans un contexte clinique : un psychologue clinicien pourrait ainsi, en intégrant l'ensemble des informations collectées pendant la passation (anamnèse, observation du patient, score total, interprétation du pattern de forces et de faiblesses sur les différents subtests, analyse par items, discussion avec le patient sur le contenu de ses réponses), analyser en profondeur quel ou quels mécanismes sont à l'origine de la performance observée - qu'elle soit particulièrement faible ou particulièrement élevée. Cette grille d'analyse pourrait servir de base à la restitution des résultats de la passation de tests dans le cadre d'un compte-rendu de bilan psychologique, et permettrait d'étayer la prise en compte de l'ensemble des déterminants de la performance pour parvenir à une véritable approche idiographique du sujet.

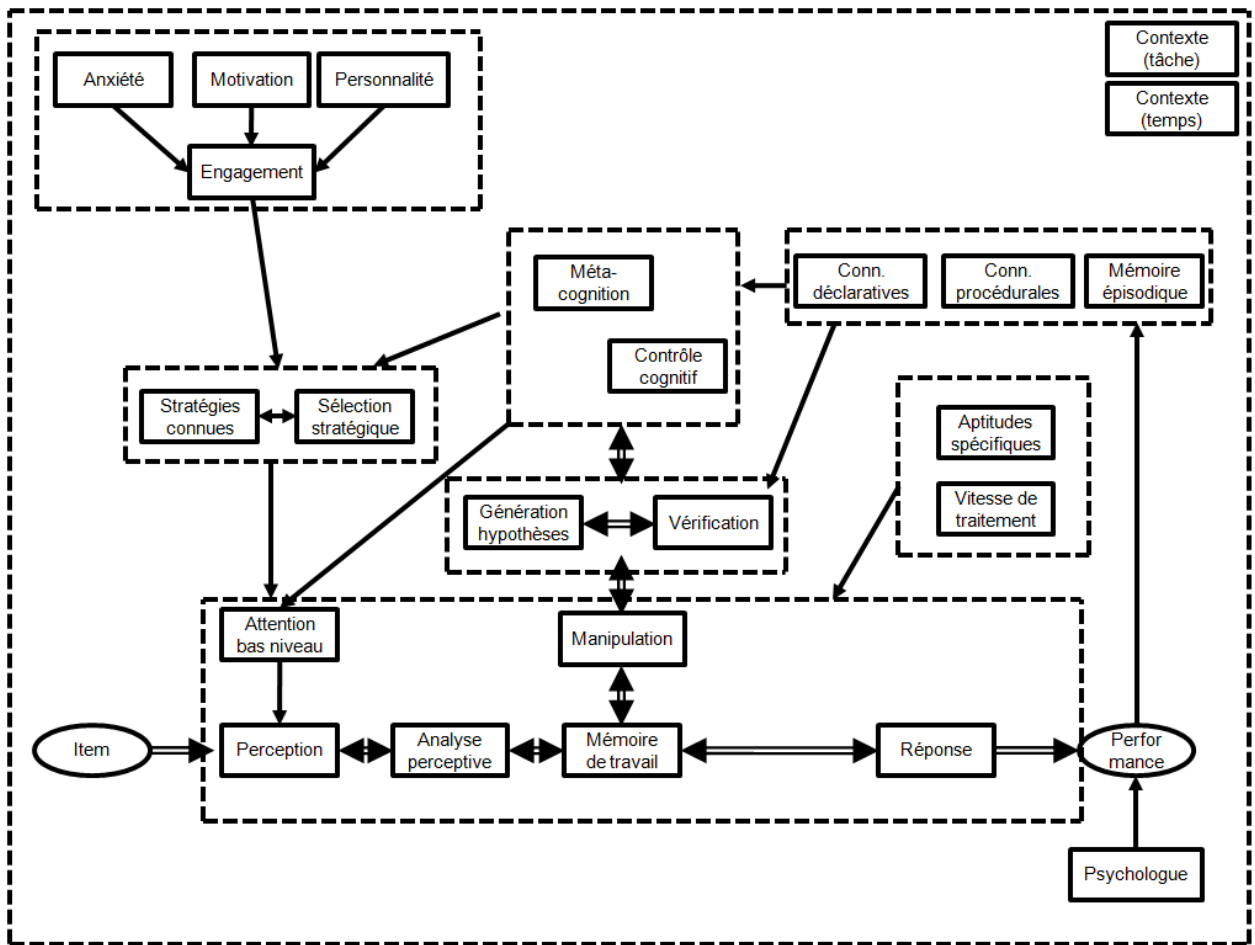


Figure 18. Ébauche d'une proposition de synthèse du système de déterminants en jeu dans les tests d'intelligence (CAR).

6. Publications et communications associées

Articles publiés et en révision

Gonthier, C., Grégoire, J., & Besançon, M. (en révision). No negative Flynn effect in France: the perils of comparing different versions of the same test.

Gonthier, C., & Roulin, J.-L. (2019). Intraindividual strategy shifts in Raven's matrices, and their dependence on working memory capacity and need for cognition. *Journal of Experimental Psychology: General*. doi:10.1037/xge0000660

Gonthier, C., & Thomassin, N. (2015). Strategy use fully mediates the relationship between working memory capacity and Raven's matrices. *Journal of Experimental Psychology: General*, 144(5), 916-924. doi: 10.1037/xge0000101

Rivollier, G., Quinton, J.-C., **Gonthier, C.**, & Smeding, A. (2020). Looking with the computer mouse: How to unveil problem-solving strategies in matrix reasoning without eye-tracking. *Behavior Research Methods*.

Articles en préparation

Gonthier, C. (en préparation). Effects of time pressure on mechanisms of performance in Raven's matrices.

Gonthier, C. (en préparation). Multiple determinants of intelligence scores: Constructing a Causal Architecture of Reasoning (CAR) for practical use.

Gonthier, C., Trégouët, G., Bac, A., & Gavornikova-Baligand (en préparation). Reasoning in students with a specific learning disability.

Tourreix, E., Besançon, M., & **Gonthier, C.** (en préparation). What does being "gifted" actually mean? The multiple roads to high IQ and the definition of giftedness.

Communications

Gonthier, C., & Roulin, J.-L. (2016, Juin – communication orale). *La mémoire de travail n'existe pas ? Une approche multi-déterministe de la cognition de haut niveau*. XXIIèmes Journées Internationales de Psychologie Différentielle, Chambéry, France.

Gonthier, C., (2016, Janvier – communication orale). *La place de la mémoire de travail dans la cognition : Multiples visages de la capacité en mémoire de travail*. Séminaire invité, CRP-CPO, Université de Picardie Jules-Verne, Amiens, France.

Gonthier, C., (2017a, Mars – communication orale). *Vers une approche multi-déterministe de la cognition de haut niveau*. Séminaire invité, LATI, Université Paris Descartes, Paris, France.

Gonthier, C. (2017b, Juin – communication orale). *Modélisation GAMM de la variabilité intra-individuelle des comportements stratégiques sur les matrices de Raven*. 6ème atelier MODEVAIIA, Marcevol, France.

Gonthier, C. (2018, Juin – communication orale). *La variabilité des comportements stratégiques sur les matrices de Raven, et ses prédicteurs*. XXIIIèmes Journées Internationales de Psychologie Différentielle, Luxembourg, Luxembourg.

SYNTHÈSE ET INDEX THÉMATIQUE

Indépendamment de l'objet étudié (contrôle cognitif, mémoire de travail ou intelligence), mes travaux peuvent se regrouper conceptuellement en grands domaines de questionnement. Je propose ci-dessous un classement (forcément arbitraire) en six questions pour faciliter la vision globale de mes projets passés et à venir.

1. Cartographie des mécanismes qualitatifs et de leurs conditions de mise en œuvre

Ces recherches portent sur l'identification des différentes stratégies ou mécanismes qualitatifs qui permettent de résoudre une tâche cognitive complexe, et des modalités selon lesquelles ces stratégies sont déployées sur une tâche.

- a) *Chapitre 1 - Gonthier & Blaye (under review) et Gonthier & Blaye (in prep)* : vérifier si les deux mécanismes qualitatifs de contrôle, réactif et proactif, peuvent être mis en jeu chez le jeune enfant lorsque la tâche permet de les recruter de façon implicite (page 44)
- b) *Chapitre 1 - Projet 1B* : montrer que les indicateurs de contrôle extraits de l'AX-CPT mélangent le mécanisme de contrôle utilisé et leur efficacité, et que la confusion entre les deux crée des corrélations factices, notamment avec la mémoire de travail (page 53)
- c) *Chapitre 1 - Projet 1C* : contrôle proactif et contrôle réactif sont-ils des mécanismes indépendants ou peuvent-ils être utilisés conjointement par le même participant dans la même tâche, voire au sein d'un même essai de l'AX-CPT ? (page 55)
- d) *Chapitre 2 - Gonthier, 2020 (Perspectives on Psychological Science)* : quelles stratégies peuvent être utilisées dans le cadre d'une tâche de mémoire de travail visuo-spatiale ? (page 92)
- e) *Chapitre 2 - Projet 2A* : quelle est la fréquence d'utilisation de ces stratégies en mémoire visuo-spatiale ? leur fréquence de co-occurrence ? leur lien avec la performance ? (page 110)
- f) *Chapitre 3 - Projet 1C* : quelles stratégies peuvent être utilisées sur les subtests les moins étudiés des échelles de Wechsler : vocabulaire, similitudes, balances, puzzles visuels ? (page 154)

2. Précision du rôle des mécanismes qualitatifs dans l'explication de la variabilité

Ces recherches portent sur l'étude du rôle des stratégies dans l'explication des différences inter et intra-individuelles de performance dans les tâches complexes, chez l'adulte ou chez l'enfant.

- a) *Chapitre 1 - Gonthier et al., 2019 (Journal of Experimental Child Psychology)* : montrer que l'amélioration des compétences de contrôle chez l'enfant s'accompagne d'une transition du contrôle réactif vers le contrôle proactif (page 41)
- b) *Chapitre 2 - Thomassin et al., 2015 (Experimental Psychology)* : montrer par une approche expérimentale-corrélacionnelle que les stratégies jouent un rôle dans le lien entre capacité en mémoire de travail et performance dans une tâche de mémorisation complexe (page 99)
- c) *Chapitre 2 - Projet 2B* : tester si les stratégies constituent une source de variance d'erreur dans la mesure de la capacité en mémoire de travail ou au contraire un contributeur important au caractère prédictif de ces mesures, en induisant directement l'usage de stratégies efficaces vs. inefficaces (page 110)
- d) *Chapitre 3 - Gonthier & Thomassin, 2015 (Journal of Experimental Psychology: General)* : montrer par une approche expérimentale-corrélacionnelle et une analyse de médiation que le lien entre capacité en mémoire de travail et performance d'intelligence peut s'expliquer par l'usage du constructive matching et de l'élimination de réponses (page 129)
- e) *Chapitre 3 - Projet 1A* : chercher à répliquer le résultat précédent, dans le contexte des résultats nuls de Jastrzębski et collaborateurs, en utilisant une mesure comportementale plus directe des stratégies utilisées (page 152)
- f) *Chapitre 3 - Gonthier & Roulin, 2019 (Journal of Experimental Psychology: General)* : montrer que le déclin de la performance au fur et à mesure de la complexification des matrices de Raven est en lien avec un déclin de l'usage du constructive matching, lui-même prédit par la mémoire de travail et le besoin de cognition (page 132)
- g) *Chapitre 3 - Projet 1B : contribution des stratégies d'intelligence à la variabilité non-adulte* : tester dans quelle mesure la variabilité de la performance d'intelligence avec l'âge (développement, vieillissement) et dans une population atypique (haut potentiel intellectuel) s'accompagne de variabilité stratégique (page 152)

3. Paradigmes pour mesurer ou manipuler les déterminants qualitatifs de la performance

Ces recherches portent sur le développement de paradigmes permettant d'évaluer quelle stratégie est mise en jeu, ou sur l'optimisation de leurs propriétés psychométriques ; ou bien sur le développement de paradigmes permettant l'induction directe de stratégies, notamment dans la perspective d'une démarche expérimentale-corrélacionnelle.

- a) *Chapitre 1 - Cooper et al., 2017 (Frontiers in Psychology)* : mettre en évidence les faibles qualités psychométriques de l'AX-CPT, la façon dont elles contribuent aux difficultés de réplication reportées dans la littérature, et la façon dont elles s'expliquent notamment par une faible variabilité des scores (page 29)
- b) *Chapitre 1 - Gonthier et al., 2016 (Frontiers in Psychology)* : développer deux paradigmes permettant d'induire l'usage respectivement du contrôle proactif et du contrôle réactif (page 48)
- c) *Chapitre 1 - Gonthier & Blaye (in prep)* : tester si l'induction d'un contrôle proactif est possible chez l'enfant d'âge préscolaire et si cette induction suffit à résorber la différence de performance avec les enfants plus âgés (page 51)
- d) *Chapitre 1 - Gonthier et al., 2016 (Memory and Cognition)* : vérifier si les effets de proportion de congruence (LWPC et ISPC) peuvent être utilisés pour mesurer l'efficacité

- du contrôle proactif et du contrôle réactif, en étudiant notamment si leur pattern de coûts et de bénéfices correspond aux prédictions (page 34)
- e) *Chapitre 1 - Projet 2B : locus des effets LWPC/ISPC* : vérifier, à l'aide du paradigme de décomposition du conflit en task conflict, semantic conflict et response conflict, à quel niveau se manifestent les effets LWPC/ISPC et si ce niveau est cohérent avec les prédictions théoriques concernant le contrôle proactif et réactif (page 57)
 - f) *Chapitre 1 - Projet 2A : réplication dans la flanker task* : vérifier si les résultats précédents peuvent se décliner à un autre type de tâche de conflit, la flanker task (page 57)
 - g) *Chapitre 1 - Bugg & Gonthier, 2020 (Quarterly Journal of Experimental Psychology)* : vérifier si l'effet LWPC apparaissant comme une manifestation du contrôle proactif peut se retrouver dans la flanker task (page 36)
 - h) *Chapitre 3 - Rivollier et al., 2020 (Behavior Research Methods)* : développer une version modifiée des matrices de Raven permettant de mesurer l'utilisation relative du constructive matching et de l'élimination de réponses en présentant alternativement soit la matrice, soit les réponses proposées (page 135)

4. Etude des processus fonctionnels qui sous-tendent la performance au niveau individuel

Ces recherches portent sur l'étude des processus qui sous-tendent la performance à un niveau de granularité individuel, c'est-à-dire au niveau de chaque essai d'une tâche, et sur la façon dont les caractéristiques de l'essai influencent ces processus de réponse.

- a) *Chapitre 1 - Tafuro et al. (under review)* : évaluer la dynamique temporelle de déploiement du contrôle cognitif au sein de chaque essai d'une tâche de conflit, en analysant les trajectoires de réponse à la souris (page 37)
- b) *Chapitre 1 - Projet 3A* : étendre le paradigme précédent à l'étude de la variabilité développementale et à d'autres paradigmes impliquant le contrôle cognitif, notamment le trail making test (page 60)
- c) *Chapitre 1 - Projet 2C* : examiner la façon dont le conflit sémantique peut générer l'interférence observée sur une tâche de conflit, en vérifiant si le conflit augmente proportionnellement au degré d'incompatibilité entre les stimuli, dans la perspective d'étudier ultérieurement l'existence d'éventuelles différences inter-individuelles (page 58)
- d) *Chapitre 2 - Gonthier & Hot, 2013 (Revue de neuropsychologie)* : mettre en évidence la complexité fonctionnelle du processus mnésique en relevant ses étapes peu étudiées par la littérature (activité pré-encodage, pré-récupération et post-récupération), telles que révélées par l'électroencéphalographie (page 83)
- e) *Chapitre 2 - Gonthier & Gavornikova-Baligand (under review)* : tester un rôle fonctionnel possible pour le contrôle attentionnel dans les tâches de mémoire de travail : contribuer au rafraîchissement des traces mnésiques en gérant l'alternance entre activités de stockage et de récupération, dans la perspective du modèle TBRS (page 88)
- f) *Chapitre 3 - Gonthier (in prep)* : étudier le rôle des temps de traitement dans les tests de raisonnement complexes, en testant de quelle façon la pression temporelle affecte les processus qui sous-tendent la réponse sur les matrices de Raven (page 139)
- g) *Chapitre 3 - Gonthier, Tréguët, Bac, & Gavornikova-Baligand (in prep)* : étudier de quelle façon l'investissement dans un test de raisonnement complexe, tel que reflété

par les temps de réponse, leur modulation en fonction de la complexité de la tâche, et l'auto-évaluation de la performance, contribue aux faibles performances observées dans les classes SEGPA (page 141)

5. Mise en évidence de la multi-détermination et de ses conséquences

Ces recherches étudient la multiplicité des processus à l'origine de la performance sur les tâches complexes, et la façon dont la reconnaissance de ce caractère multi-déterminé peut modifier l'interprétation de la performance (par rapport à une posture réaliste).

- a) *Chapitre 1 - Projet 3B* : étudier les déterminants possibles de la transition développementale du contrôle réactif vers le contrôle proactif (comme une fenêtre possible vers la compréhension plus générale des déterminants du contrôle), notamment le rôle de la métacognition et de la connaissance ou non des stratégies possibles (page 61)
- b) *Chapitre 2 - Bagneux et al., 2013 (PLoS ONE)* : étudier le rôle de la capacité en mémoire de travail dans l'IGT, une tâche de prise de décision traditionnellement considérée comme étant fondée sur un apprentissage implicite (page 95)
- c) *Chapitre 2 - Aubry et al. (under review)* : vérifier si la capacité en mémoire de travail particulièrement élevée des enfants à haut potentiel est en lien avec un contrôle cognitif plus efficace que la norme, ou si ce sont les aptitudes tâches-spécifiques qui en sont à l'origine (page 91)
- d) *Chapitre 3 - Gonthier, Grégoire, & Besançon (under review)* : montrer que le soi-disant déclin de l'intelligence en France reflète en réalité une évolution des connaissances déclaratives dans le contexte de changements culturels, à l'origine d'une performance plus faible sur les items anciens (page 146)
- e) *Chapitre 3 - Toureix, Besançon, & Gonthier (in prep)* : vérifier s'il existe plusieurs profils de personnes à haut potentiel sur les échelles de Wechsler, qui correspondraient à une supériorité sur différents déterminants de la performance ; avec en particulier la possibilité d'une surestimation de la performance sur les tests évaluant les connaissances déclaratives (page 148)
- f) *Chapitre 3 - Projet 2B* : tester un effet possible des connaissances procédurales en vérifiant s'il existe un effet du sens de lecture de la langue maternelle (gauche-droite vs. droite-gauche) sur une épreuve de type matrices de Raven (page 155)
- g) *Chapitre 3 - Projet 3* : réaliser une synthèse globale de l'architecture causale à l'origine de la performance sur un test d'intelligence, présentée sous la forme d'un système de déterminants quantitatifs et qualitatifs inscrits dans un contexte donné (projet CAR : page 157)

6. Prise en compte de multiples déterminants dans l'évaluation de la performance

Ces recherches portent sur le développement d'épreuves permettant d'évaluer la cognition de haut niveau, en tenant compte de la multiplicité des déterminants à l'origine de la performance, dans le but de minimiser les sources de variance non pertinentes ; ces projets ont

à la fois un objectif d'utilisation en recherche fondamentale, et la perspective à long terme de développer de meilleures évaluations cliniques.

- a) *Chapitre 1 - Projet 1A* : réaliser une synthèse de la multiplicité des versions du paradigme AX-CPT, en listant toutes les variantes qui peuvent interagir avec certains déterminants de la performance (limite de temps, présence de feedbacks, délimitation indice/cible, etc), de façon à proposer une version de référence (page 52)
- b) *Chapitre 2 - Gonthier et al., 2016 (Behavior Research Methods)* : création d'une tâche composite de mémoire de travail, le CCS, intégrant trois types d'empans complexes pour obtenir une mesure domaine-générale brève (page 75)
- c) *Chapitre 2 - Projet 1B* : étendre la logique de l'extraction d'une mesure domaine-générale de la capacité en mémoire de travail en combinant des tâches de mémorisation intégrant non seulement plusieurs modalités sensorielles, mais aussi plusieurs paradigmes différents (page 102)
- d) *Chapitre 2 - Projet 1C* : améliorer la précision du scoring de tâches de mémoire de travail en expérimentant un nouveau mode de comptabilisation de la performance, notamment la distance d'édition, et en vérifiant s'il est possible d'étendre les tâches de mémorisation visuo-spatiales à des espaces continus (page 106)
- e) *Chapitre 2 - Gonthier et al., 2017 (Behavior Research Methods)* : extension de la tâche composite du CCS à un paradigme adaptatif pour créer une tâche, l'ACCES, qui permet l'évaluation de populations qui présentent des niveaux d'aptitude très différents (page 81)
- f) *Chapitre 2 - Projet 1A* : comparer une version classique et une version adaptative de la même épreuve pour déterminer dans quelle mesure le caractère adaptatif affecte les propriétés psychométriques, mais aussi les aspects motivationnels tels que la fatigue et l'engagement dans la tâche (page 102)
- g) *Chapitre 2 - Projet 1D* : s'appuyer sur les travaux précédents pour développer une épreuve de mémoire de travail à vocation clinique, permettant d'évaluer des populations d'aptitude faible, et surtout, qui présentent un déficit important dans les tâches de jugement simples qui forment la base des empans complexes (page 107)
- h) *Chapitre 2 - Projet 1E* : vérifier dans quelle mesure les paradigmes de double tâche utilisés notamment en neuropsychologie clinique permettent réellement d'évaluer les différences inter-individuelles de contrôle cognitif et/ou mémoire de travail, en cherchant à montrer que ces paradigmes peuvent produire un hard fall effect qui inverse l'interprétation requise pour les scores de chute de performance de type mu (page 108)
- i) *Chapitre 3 - Projet 2A* : réaliser une tâche adaptative de raisonnement, sur le modèle de la tâche adaptative de mémoire de travail ACCES, qui faciliterait l'évaluation et la comparaison de populations présentant des niveaux d'aptitude très différents, que ce soit dans un contexte développemental, ou atypique tel que le HPI (page 155)
- j) *Chapitre 3 - Projet 3* : s'appuyer sur un modèle systémique de la multi-détermination des scores d'intelligence (CAR) pour décliner ce modèle sous la forme d'une grille d'aide à l'interprétation des scores d'intelligence, tels que collectés dans le cadre d'une épreuve de type Wechsler (page 162).

BIBLIOGRAPHIE

- Ackerman, P. L. (2016). Process overlap and g do not adequately account for a general factor of intelligence. *Psychological Inquiry*, 27(3), 178–180. doi:10.1080/1047840X.2016.1158087
- Ackerman, P. L., Beier, M. E., & Boyle, M. D. (2002). Individual differences in working memory within a nomological network of cognitive and perceptual speed abilities. *Journal of Experimental Psychology: General*, 131(4), 567-589. doi:10.1037/0096-3445.131.4.567
- Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working Memory and Intelligence: The Same or Different Constructs? *Psychological Bulletin*, 131(1), 30-60. doi:10.1037/0033-2909.131.1.30
- Ackerman, P. L., & Cianciolo, A. T. (2002). Ability and task constraint determinants of complex task performance. *Journal of Experimental Psychology: Applied*, 8(3), 194–208. doi:10.1037/1076-898X.8.3.194
- Adam, K. C. S., Mance, I., Fukuda, K., & Vogel, E. K. (2015). The contribution of attentional lapses to individual differences in visual working memory capacity. *Journal of Cognitive Neuroscience*, 27(8), 1601-1616. doi:10.1162/jocn_a_00811
- Adams, E. J., Nguyen, A. T., & Cowan, N. (2018). Theories of working memory: Differences in definition, degree of modularity, role of attention, and purpose. *Language, Speech, and Hearing Services in Schools*, 49(3), 340–355. doi:10.1044/2018_LSHSS-17-0114
- Ali, A., Ambler, G., Strydom, A., Rai, D., Cooper, C., McManus, S., Weich, S., Meltzer, H., Dein, S., & Hassiotis, A. (2013). The relationship between happiness and intelligent quotient: The contribution of socio-economic and clinical factors. *Psychological Medicine*, 43(6), 1303–1312. doi:10.1017/S0033291712002139
- Allan, K., & Rugg, M. D. (1997). An event-related potential study of explicit memory on tests of cued recall and recognition. *Neuropsychologia*, 35(4), 387–397. doi:10.1016/S0028-3932(96)00094-2
- Alloway, T. P. (2007). *Automated Working Memory Assessment (AWMA)*. Harcourt Assessment.
- Alloway, T. P., Gathercole, S. E., & Elliott, J. (2010). Examining the link between working memory behaviour and academic attainment in children with ADHD. *Developmental Medicine & Child Neurology*, 52(7), 632–636. doi:10.1111/j.1469-8749.2009.03603.x

- Alloway, R. G., Alloway, T. P., Magyari, P. M., & Floyd, S. (2016). An exploratory study investigating the effects of barefoot running on working memory. *Perceptual and Motor Skills, 122*(2), 432–443. doi:10.1177/0031512516640391
- Anderson, M. L. (2010). Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain sciences, 33*, 245–313. doi:10.1017/S0140525X10000853
- Anderson, V., Jacobs, R., & Anderson, P. J. (2008). *Executive functions and the frontal lobes: A lifespan perspective*. Taylor & Francis.
- Ang, S. Y., & Lee, K. (2010). Exploring developmental differences in visual short-term memory and working memory. *Developmental Psychology, 46*(1), 279–285. doi:10.1037/a0017554
- Ardiale, E., & Lemaire, P. (2013). Within-item strategy switching in arithmetic: A comparative study in children. *Frontiers in Psychology, 4*. doi:10.3389/fpsyg.2013.00924
- Arendasy, M. E., & Sommer, M. (2013). Reducing response elimination strategies enhances the construct validity of figural matrices. *Intelligence, 41*(4), 234–243. doi:doi.org/10.1016/j.intell.2013.03.006
- Arsalidou, M. (2013). Working memory capacity: The need for process task-analysis. *Frontiers in Psychology, 4*. doi:10.3389/fpsyg.2013.00257
- Arvey, R. D. (1972). Some comments on culture fair tests. *Personnel Psychology, 25*(3), 433–448. doi:10.1111/j.1744-6570.1972.tb00828.x
- Ashton, M. C., Lee, K., Goldberg, L. R., & de Vries, R. E. (2009). Higher order factors of personality: Do they exist? *Personality and Social Psychology Review, 13*(2), 79–91. doi:10.1177/1088868309338467
- Atkins, S. M., Sprenger, A. M., Colflesh, G. J. H., Briner, T. L., Buchanan, J. B., Chavis, S. E., Chen, S., Iannuzzi, G. L., Kashtelyan, V., Dowling, E., Harbison, J. I., Bolger, D. J., Bunting, M. F., & Dougherty, M. R. (2014). Measuring working memory is all fun and games: A four-dimensional spatial game predicts cognitive task performance. *Experimental Psychology, 61*(6), 417–438. doi:10.1027/1618-3169/a000262
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence and J. T. Spence (Eds.), *Psychology of learning and motivation, 2* (pp. 89–195). Oxford, UK: Academic Press. doi:10.1016/S0079-7421(08)60422-3
- Aubry, A., Gonthier, C., & Bourdin, B. (en révision). Explaining the high working memory capacity of gifted children: Contributions of processing skills and executive control. *High Ability Studies*.

- Augustinova, M., Silvert, L., Spatola, N., & Ferrand, L. (2018). Further investigation of distinct components of Stroop interference and of their reduction by short response-stimulus intervals. *Acta Psychologica, 189*, 54–62. doi:10.1016/j.actpsy.2017.03.009
- Bacon, A. M., & Handley, S. J. (2010). Dyslexia and reasoning: The importance of visual processes. *British Journal of Psychology, 101*(3), 433–452. doi:10.1348/000712609X467314
- Bacon, A. M., Handley, S. J., & McDonald, E. L. (2007). Reasoning and dyslexia: A spatial strategy may impede reasoning with visually rich information. *British Journal of Psychology, 98*(1), 79–92. doi:10.1348/000712606X103987
- Baddeley, A. D. (1986). *Working Memory*. Oxford: Clarendon Press.
- Baddeley, A. D. (1996). Exploring the Central Executive. *The Quarterly Journal of Experimental Psychology: Section A, 49*(1), 5–28. doi:10.1080/027249896392784
- Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends In Cognitive Sciences, 4*(11), 417–423. doi:10.1016/S1364-6613(00)01538-2
- Baddeley, A. D. (2002). Fractionating the central executive. In D. T. Stuss & R. T. Knight (Eds.), *Principles of Frontal Lobe Function* (pp. 246–260). Oxford University Press.
- Baddeley, A. D. (2003). Working Memory: Looking Back and Looking Forward. *Nature Reviews Neuroscience, 4*(10), 829–839. doi:10.1038/nrn1201
- Baddeley, A., Della Sala, S., Papagno, C., & Spinnler, H. (1997). Dual-task performance in dysexecutive and nondysexecutive patients with a frontal lesion. *Neuropsychology, 11*(2), 187–194. doi:10.1037/0894-4105.11.2.187
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47–89). Academic Press.
- Bagneux, V., Thomassin, N., Gonthier, C., & Roulin, J.-L. (2013). Working memory in the processing of the Iowa Gambling Task: An individual differences approach. *PLoS ONE, 8*(11) : e81498. doi : 10.1371/journal.pone.0081498
- Bailey, H., Dunlosky, J., & Kane, M. J. (2008). Why does working memory span predict complex cognition? Testing the strategy affordance hypothesis. *Memory & Cognition, 36*(8), 1383–1390. doi:10.3758/MC.36.8.1383
- Barch, D. M., Carter, C. S., MacDonald, A. W., III, Braver, T. S., & Cohen, J. D. (2003). Context-processing deficits in schizophrenia: Diagnostic specificity, 4-week course, and relationships to clinical symptoms. *Journal of Abnormal Psychology, 112*(1), 132–143. doi:10.1037/0021-843X.112.1.132

- Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, *121*(1), 65-94. doi:10.1037/0033-2909.121.1.65
- Barrouillet, P., & Camos, V. (2007). The time-based resource sharing model of working memory. In N. Osaka, R. H. Logie & M. D'Esposito (Eds.), *The cognitive neuroscience of working memory* (pp. 59-80). New York, NY: Oxford University Press.
- Barrouillet, P., Gavens, N., Vergauwe, E., Gaillard, V., & Camos, V. (2009). Working memory span development: A time-based resource-sharing model account. *Developmental Psychology*, *45*(2), 477-490. doi:10.1037/a0014615
- Barrouillet, P., Lépine, R., & Camos, V. (2008). Is the influence of working memory capacity on high-level cognition mediated by complexity or resource-dependent elementary processes? *Psychonomic Bulletin & Review*, *15*(3), 528-534. doi:10.3758/PBR.15.3.528
- Bartholomew, D. J., Allerhand, M., & Deary, I. J. (2013). Measuring mental capacity: Thomson's Bonds model and Spearman's g-model compared. *Intelligence*, *41*(4), 222-233. doi:10.1016/j.intell.2013.03.007
- Bartholomew, D. J., Deary, I. J., & Lawn, M. (2009). A new lease of life for Thomson's bonds model of intelligence. *Psychological Review*, *116*(3), 567-579. doi:10.1037/a0016262
- Batty, D., Wennerstad, K. M., Smith, G. D., Gunnell, D., Deary, I. J., Tynelius, P., & Rasmussen, F. (2009). IQ in early adulthood and mortality by middle age: Cohort study of 1 million Swedish men. *Epidemiology*, *20*(1), 100-109. doi:10.1097/EDE.0b013e31818ba076
- Baxendale, S. (2011). IQ and ability across the adult life span. *Applied Neuropsychology*, *18*(3), 164-167. doi:10.1080/09084282.2011.595442
- Beaujean, A., & Sheng, Y. (2010). Examining the Flynn effect in the general social survey vocabulary test using item response theory. *Personality and Individual Differences*, *48*(3), 294-298. doi:10.1016/j.paid.2009.10.019
- Bechara, A., Damasio, H., Tranel, D., & Anderson, S. W. (1998). Dissociation of working memory from decision making within the human prefrontal cortex. *The Journal of Neuroscience*, *18*(1), 428-437.
- Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, *275*(5304), 1293-1295. doi:10.1126/science.275.5304.1293
- Becker, N. Schmitz, F. Falk, A. M. Feldbrügge, J. Recktenwald, D. R. Wilhelm, O. Preckel, F. & Spinath, F. M. (2016). Preventing response elimination strategies improves the convergent validity of figural matrices. *Journal of Intelligence*, *4*(2). doi:10.3390/jintelligence4010002

- Becker, N., Schmitz, F., Göritz, A. S., & Spinath, F. M. (2016). Sometimes more is better, and sometimes less is better: Task complexity moderates the response time accuracy correlation. *Journal of Intelligence*, 4(11). doi:10.3390/jintelligence4030011
- Beilock, S. L., & Carr, T. H. (2005). When High-Powered People Fail: Working Memory and 'Choking Under Pressure' in Math. *Psychological Science*, 16(2), 101-105. doi:10.1111/j.0956-7976.2005.00789.x
- Beilock, S. L., & DeCaro, M. S. (2007). From poor performance to success under stress: Working memory, strategy selection, and mathematical problem solving under pressure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(6), 983-998. doi:10.1037/0278-7393.33.6.983
- Benton, S. L., Kraft, R. G., Glover, J. A., & Plake, B. S. (1984). Cognitive capacity differences among writers. *Journal Of Educational Psychology*, 76(5), 820-834. doi:10.1037/0022-0663.76.5.820
- Bethell-Fox, C. E., Lohman, D. F., & Snow, R. E. (1984). Adaptive reasoning: Componential and eye movement analysis of geometric analogy performance. *Intelligence*, 8(3), 205–238. doi:10.1016/0160-2896(84)90009-6
- Blair, C. (2006). How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. *Behavioral And Brain Sciences*, 29(2), 109-160. doi:10.1017/S0140525X06009034
- Blais, C., Harris, M. B., Guerrero, J. V., & Bunge, S. A. (2012). Rethinking the role of automaticity in cognitive control. *The Quarterly Journal of Experimental Psychology*, 65(2), 268–276. doi:10.1080/17470211003775234
- Blum, D., & Holling, H. (2017). Spearman's law of diminishing returns: A meta-analysis. *Intelligence*, 65, 60–66. doi:10.1016/j.intell.2017.07.004
- Bock, G. R., Goode, J. A., & Webb, K. (2000). *The Nature of Intelligence: Novartis Foundation Symposium* (Vol. 233). John Wiley & Sons.
- Bock, O., Haeger, M., & Voelcker-Rehage, C. (2019). Structure of executive functions in young and in older persons. *Plos ONE*, 14(5):e0216149. doi:10.1371/journal.pone.0216149
- Borsboom, D., Kievit, R. A., Cervone, D., & Hood, S. B. (2009). The two disciplines of scientific psychology, or: The disunity of psychology as a working hypothesis. In J. Valsiner, P. C. M. Molenaar, M. C. D. P. Lyra, & N. Chaudhary (Eds.), *Dynamic process methodology in the social and developmental sciences*. (pp. 67–97). Springer Science + Business Media. doi:10.1007/978-0-387-95922-1_4

- Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110*(2), 203–219. doi:10.1037/0033-295X.110.2.203
- Braem, S., Bugg, J. M., Schmidt, J. R., Crump, M. J. C., Weissman, D. H., Notebaert, W., & Egner, T. (2019). Measuring adaptive control in conflict tasks. *Trends in Cognitive Sciences*, *23*(9), 769–783. doi:10.1016/j.tics.2019.07.002
- Brand, C. R., Freshwater, S., & Dockrell, W. B. (1989). Has there been a “massive” rise in IQ levels in the West? Evidence from Scottish children. *The Irish Journal of Psychology*, *10*(3), 388–393. doi:10.1080/03033910.1989.10557756
- Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences*, *16*(2), 106–113. doi:10.1016/j.tics.2011.12.010
- Braver, T. S., Barch, D. M., Keys, B. A., Carter, C. S., Cohen, J. D., Kaye, J. A., Janowsky, J. S., Taylor, S. F., Yesavage, J. A., Mumenthaler, M. S., Jagust, W. J., & Reed, B. R. (2001). Context processing in older adults: Evidence for a theory relating cognitive control to neurobiology in healthy aging. *Journal of Experimental Psychology: General*, *130*(4), 746–763. doi:10.1037/0096-3445.130.4.746
- Braver, T. S., Gray, J. R., & Burgess, G. C. (2007). Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. In A. R. A. Conway, C. Jarrold, M. J. Kane, J. N. Towse, & A. Miyake (Eds.), *Variation in working memory* (pp. 76-106). New York, NY, US: Oxford University Press.
- Braver, T. S., Paxton, J. L., Locke, H. S., & Barch, D. M. (2009). Flexible neural mechanisms of cognitive control within human prefrontal cortex. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *106*(18), 7351-7356. doi:10.1073/pnas.08081871067
- Braver, T. S., Satpute, A. B., Rush, B. K., Racine, C. A., & Barch, D. M. (2005). Context Processing and Context Maintenance in Healthy Aging and Early Stage Dementia of the Alzheimer’s Type. *Psychology and Aging*, *20*(1), 33–46. doi:10.1037/0882-7974.20.1.33
- Broadbent, D. E. (1984). The maltese cross: A new simplistic model for memory. *Behavioral And Brain Sciences*, *7*(1), 55-94. doi:10.1017/S0140525X00026121
- Broadway, J. M., & Engle, R. W. (2010). Validating running memory span: Measurement of working memory capacity and links with fluid intelligence. *Behavior Research Methods*, *42*(2), 563–570. doi:10.3758/BRM.42.2.563
- Brose, A., Schmiedek, F., Lövdén, M., & Lindenberger, U. (2012). Daily variability in working memory is coupled with negative affect: The role of attention and motivation. *Emotion*, *12*(3), 605-617. doi:10.1037/a0024436

- Bugg, J. M. (2015). The relative attractiveness of distractors and targets affects the coming and going of item-specific control: Evidence from flanker tasks. *Attention, Perception, & Psychophysics*, *77*(2), 373–389. doi:10.3758/s13414-014-0752-x
- Bugg, J., & Gonthier, C. (2020). List-level control in the flanker task. *Quarterly Journal of Experimental Psychology*, *73*(9), 1444–1459. doi:10.1177/1747021820912477
- Bugg, J. M., Jacoby, L. L., & Chanani, S. (2011). Why it is too early to lose control in accounts of item-specific proportion congruency effects. *Journal of Experimental Psychology: Human Perception And Performance*, *37*(3), 844 – 859. doi:10.1037/a0019957
- Bugg, J. M., Jacoby, L. L., & Toth, J. P. (2008). Multiple levels of control in the Stroop task. *Memory & Cognition*, *36*(8), 1484–1494. doi:10.3758/MC.36.8.1484
- Buehner, M., Krumm, S., Ziegler, M., & Pluecken, T. (2006). Cognitive abilities and their interplay: Reasoning, crystallized intelligence, working memory components, and sustained attention. *Journal of Individual Differences*, *27*(2), 57–72. doi:10.1027/1614-0001.27.2.57
- Burt, C. (1940). *The factors of the mind*. University of London Press.
- Buszard, T., Farrow, D., Verswijveren, S. J. J. M., Reid, M., Williams, J., Polman, R., Ling, F. C. M., & Masters, R. S. W. (2017). Working memory capacity limits motor learning when implementing multiple instructions. *Frontiers in Psychology*, *8*. doi:10.3389/fpsyg.2017.01350
- Camos, V., Lagner, P., & Barrouillet, P. (2009). Two maintenance mechanisms of verbal information in working memory. *Journal of Memory and Language*, *61*(3), 457–469. doi:10.1016/j.jml.2009.06.002
- Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, *97*(3), 404–431. doi:10.1037/0033-295x.97.3.404
- Carroll, J. B. (1991). No demonstration that g is not unitary, but there's more to the story: Comment on Kranzler and Jensen. *Intelligence*, *15*(4), 423–436. doi:10.1016/0160-2896(91)90004-W
- Cary, M., & Reder, L. M. (2002). Metacognition in strategy selection. In P. Chambres, M. Izaute, & P.-J. Marescaux (Eds.), *Metacognition: Process, function and use* (pp. 63–77). Dordrecht: Kluwer Academic Publishers. doi:10.1007/978-1-4615-1099-4_5
- Case, R., Kurland, M., & Daneman, M. (1979, March). *Operational efficiency and the growth of M-space* [Paper presentation]. Society for Research in Child Development, San Francisco.

- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54*(1), 1–22. doi:10.1037/h0046743
- Chatham, C. H., Frank, M. J., & Munakata, Y. (2009). Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control. *PNAS Proceedings Of The National Academy Of Sciences Of The United States Of America, 106*(14), 5529-5533. doi:10.1073/pnas.0810002106
- Chein, J. M., Moore, A. B., & Conway, A. R. A. (2011). Domain-general mechanisms of complex working memory span. *Neuroimage, 54*(1), 550-559. doi:10.1016/j.neuroimage.2010.07.067
- Chevalier, N. (2015). The development of executive function: Toward more optimal coordination of control with age. *Child Development Perspectives, 9*(4), 239-244. doi:10.1111/cdep.12138
- Chevalier, N., Martis, S. B., Curran, T., & Munakata, Y. (2015). Metacognitive processes in executive control development: The case of reactive and proactive control. *Journal Of Cognitive Neuroscience, 27*(6), 1125-1136. doi:10.1162/jocn_a_00782
- Cohen-Shikora, E. R., Diede, N. T., & Bugg, J. M. (2018). The flexibility of cognitive control: Age equivalence with experience guiding the way. *Psychology and Aging, 33*(6), 924–939. doi:10.1037/pag00002807
- Cokely, E. T., Kelley, C. M., & Gilchrist, A. L. (2006). Sources of individual differences in working memory: Contributions of strategy to capacity. *Psychonomic Bulletin & Review, 13*(6), 991-997. doi:10.3758/BF03213914
- Collette, F., & van der Linden, M. (2002). Brain imaging of the central executive component of working memory. *Neuroscience and Biobehavioral Reviews, 26*(2), 105–125. doi:10.1016/S0149-7634(01)00063-X
- Colom, R., Rebollo, I., Abad, F. J., & Shih, P. C. (2006). Complex span tasks, simple span tasks, and cognitive abilities: A reanalysis of key studies. *Memory & Cognition, 34*(1), 158–171. doi:10.3758/BF03193395
- Conway, A. R. A., Cowan, N., Bunting, M. F., Theriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence, 30*(2), 163-183. doi:10.1016/S0160-2896(01)00096-4
- Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks : A methodological review and user's guide. *Psychonomic bulletin & review, 12*(5), 769-786. doi:10.3758/BF03196772

- Cooper, S. R., Gonthier, C., Barch, D. M., & Braver, T. S. (2017). The role of psychometrics in individual differences research in cognition: A case study of the AX-CPT. *Frontiers in Psychology, 8*(1482). doi:10.3389/fpsyg.2017.01482
- Copeland, D. E., & Radvansky, G., (2004). Working memory and syllogistic reasoning. *The Quarterly Journal Of Experimental Psychology A: Human Experimental Psychology, 57A*(8), 1437-1457. doi:10.1080/02724980343000846
- Cornoldi, C., & Vecchi, T. (2003). *Visuo-spatial working memory and individual differences*. Hove: Psychology Press.
- Corroyer, D., & Rozencajg, P. (1995). "Samuel", un outil de diagnostic automatique des stratégies dans la tâche des cubes de Kohs. Delta-Expert. Available at: <http://www.delta-expert.com>
- Cowan, N. (1999). An Embedded-Processes Model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge University Press. doi:10.1017/CBO9781139174909.006
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral And Brain Sciences, 24*(1), 87-185. doi:10.1017/S0140525X01003922
- Cowan, N., Saults, J. S., & Morey, C. C. (2006). Development of working memory for verbal-spatial associations. *Journal of Memory and Language, 55*(2), 274–289. doi:10.1016/j.jml.2006.04.002
- Cowan, N., Wood, N. L., Wood, P. K., Keller, T. A., Nugent, L. D., & Keller, C. V. (1998). Two separate verbal processing rates contributing to short-term memory span. *Journal of Experimental Psychology: General, 127*(2), 141–160. doi:10.1037/0096-3445.127.2.141
- Coyle, T. R., Read, L. E., Gaultney, J. F., & Bjorklund, D. F. (1998). Giftedness and variability in strategic processing on a multitrial memory task: Evidence for stability in gifted cognition. *Learning and Individual Differences, 10*(4), 273–290. [http://doi.org/10.1016/S1041-6080\(99\)80123-X](http://doi.org/10.1016/S1041-6080(99)80123-X)
- Craik, F. I. M., Bialystok, E., Gillingham, S., & Stuss, D. T. (2018). Alpha span: A measure of working memory. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 72*(3), 141–152. doi:10.1037/cep0000143
- Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist, 12*(11), 671-684. doi:10.1037/h0043943
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*(4), 281–302. doi:10.1037/h0040957

- Crump, M. J. C., Gong, Z., & Milliken, B. (2006). The context-specific proportion congruent Stroop effect: Location as a contextual cue. *Psychonomic Bulletin & Review*, *13*(2), 316–321. doi:10.3758/BF03193850
- Crump, M. J. C., Vaquero, J. M. M., & Milliken, B. (2008). Context-specific learning and control: The roles of awareness, task relevance, and relative salience. *Consciousness and Cognition: An International Journal*, *17*(1), 22–36. doi:10.1016/j.concog.2007.01.004
- Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal Of Experimental Psychology: Learning, Memory, And Cognition*, *9*(4), 561-584. doi:10.1037/0278-7393.9.4.561
- Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*(4), 422-433. doi:10.3758/BF03214546
- Danthiir, V., Roberts, R. D., Schulze, R., & Wilhelm, O. (2005). Mental Speed: On Frameworks, Paradigms, and a Platform for the Future. In O. Wilhelm & R. W. Engle (Eds.), *Handbook of understanding and measuring intelligence* (pp. 27-46). Thousand Oaks, CA, US: Sage Publications, Inc. doi:10.4135/9781452233529.n3
- De Dreu, C. K. W., Nijstad, B. A., Baas, M., Wolsink, I., & Roskes, M. (2012). Working memory benefits creative insight, musical improvisation, and original ideation through maintained task-focused attention. *Personality and Social Psychology Bulletin*, *38*(5), 656–669. doi:10.1177/0146167211435795
- De Marchis, G., del Prado Rivero Expósito, M., & Avilés, J. M. R. (2013). Psychological distance and reaction time in a Stroop task. *Cognitive Processing*, *14*(4), 401–410. doi:10.1007/s10339-013-0569-x
- Deary, I. J., Lawn, M., & Bartholomew, D. J. (2008). A conversation between Charles Spearman, Godfrey Thomson, and Edward L Thorndike: The International Examinations Inquiry Meetings 1931-1938. *History of Psychology*, *11*(2), 122–142. doi:10.1037/1093-4510.11.2.122
- Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*(1), 13–21. doi:10.1016/j.intell.2006.02.001
- Delaloye, C. C., Ludwig, C. C., Borella, E. E., Chicherio, C. C., & de Ribaupierre, A. A. (2008). L'Empan de lecture comme épreuve mesurant la capacité de mémoire de travail: Normes basées sur une population francophone de 775 adultes jeunes et âgés. [The Reading Span as a measure of working memory capacity: Norms based on a French speaking population of 775 younger and older adults]. *European Review Of Applied Psychology*, *58*(2), 89-103. doi:10.1016/j.erap.2006.12.004

- Delaney, P. F., & Sahakyan, L. (2007). Unexpected costs of high working memory capacity following directed forgetting and contextual change manipulations. *Memory & Cognition*, *35*(5), 1074–1082. doi:10.3758/BF03193479
- Della Sala, S., Foley, J. A., Beschin, N., Allerhand, M., & Logie, R. H. (2010). Assessing dual-task performance using a paper-and-pencil test: Normative data. *Archives of Clinical Neuropsychology*, *25*(5), 410–419. doi:10.1093/arclin/acq039
- Desmette, D., Hupet, M., Schelstraete, M., & van der Linden, M. (1995). Adaptation en langue française du 'Reading Span Test' de Daneman et Carpenter (1980). [A French version of Daneman and Carpenter's (1980) Reading Span Test]. *L'année Psychologique*, *95*(3), 459-482. doi:10.3406/psy.1995.28842
- Detterman, D. K. (1982). Does “g” exist? *Intelligence*, *6*(2), 99–108. doi:10.1016/0160-2896(82)90008-3
- Detterman, D. K. (1987). Theoretical notions of intelligence and mental retardation. *American Journal of Mental Deficiency*, *92*(1), 2–11.
- Detterman, D. K., Petersen, E., & Frey, M. C. (2016). Process overlap and system theory: A simulation of, comment on, and integration of Kovacs and Conway. *Psychological Inquiry*, *27*(3), 200–204. doi:10.1080/1047840X.2016.1181514
- Detterman, D. K. (2000). General intelligence and the definition of phenotypes. In G. R. Bock, J. A. Goode, & K. Webb (Eds.), *The Nature of Intelligence: Novartis Foundation Symposium* (Vol. 233). John Wiley & Sons.
- Detterman, D. K., & Daniel, M. H. (1989). Correlations of mental tests with each other and with cognitive variables are highest for low IQ groups. *Intelligence*, *13*(4), 349–359. doi:10.1016/S0160-2896(89)80007-8
- DiTrapani, J., Jeon, M., De Boeck, P., & Partchev, I. (2016). Attempting to differentiate fast and slow intelligence: Using generalized item response trees to examine the role of speed on intelligence tests. *Intelligence*, *56*, 82–92. doi:10.1016/j.intell.2016.02.012
- Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64*(1), 135-168. doi:10.1146/annurev-psych-113011-143750
- Doebel, S., Dickerson, J. P., Hoover, J. D., & Munakata, Y. (2018). Using language to get ready: Familiar labels help children engage proactive control. *Journal of Experimental Child Psychology*, *166*, 147–159. doi:10.1016/j.jecp.2017.08.006
- Dori, G. A., & Chelune, G. J. (2004). Education-Stratified Base-Rate Information on Discrepancy Scores Within and Between the Wechsler Adult Intelligence Scale-Third Edition and the Wechsler Memory Scale-Third Edition. *Psychological Assessment*, *16*(2), 146–154. doi:10.1037/1040-3590.16.2.146

- Duan, X., Shi, J., Wu, J., Mou, Y., Cui, H., & Wang, G. (2009). Electrophysiological correlates for response inhibition in intellectually gifted children: A Go/NoGo study. *Neuroscience Letters*, *457*(1), 45–48. doi:10.1016/j.neulet.2009.04.006
- Duckworth, A., Quinn, P. D., Lynam, D. R., Loeber, R., & Stouthamer-Loeber, M. (2011). Role of test motivation in intelligence testing. *PNAS Proceedings Of The National Academy Of Sciences Of The United States Of America*, *108*(19), 7716–7720. doi:10.1073/pnas.1018601108
- Dunlosky, J., & Kane, M. J. (2007). The contributions of strategy use to working memory span: A comparison of strategy assessment methods. *The Quarterly Journal of Experimental Psychology*, *60*(9), 1227–1245. doi:10.1080/17470210600926075
- Dutton, E., & Lynn, R. (2015). A negative Flynn Effect in France, 1999 to 2008–9. *Intelligence*, *51*, 67–70. doi:10.1016/j.intell.2015.05.005
- Dutton, E., van der Linden, D., & Lynn, R. (2016). The negative Flynn Effect: A systematic literature review. *Intelligence*, *59*, 163–169. doi:10.1016/j.intell.2016.10.002
- Edwards, B. G., Barch, D. M., & Braver, T. S. (2010). Improving prefrontal cortex function in schizophrenia through focused training of cognitive control. *Frontiers in Human Neuroscience*, *4*, . doi:10.3389/fnhum.2010.00032
- Egan, D. E., & Grimes-Farrow, D. D. (1982). Differences in mental representations spontaneously adopted for reasoning. *Memory & Cognition*, *10*(4), 297–307. doi:10.3758/BF03202421
- Elke, S., & Wiebe, S. A. (2017). Proactive control in early and middle childhood: An ERP study. *Developmental Cognitive Neuroscience*, *26*, 28–38. doi:10.1016/j.dcn.2017.04.005
- Ellis, J. L., & Van den Wollenberg, A. L. (1993). Local homogeneity in latent trait models: A characterization of the homogeneous monotone IRT model. *Psychometrika*, *58*(3), 417–429. doi:10.1007/BF02294649
- Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. *Psychology of Learning & Motivation*(44), 145.
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*(3), 309–331. doi:10.1037/0096-3445.128.3.309
- Erb, C. D., Moher, J., Sobel, D. M., & Song, J.-H. (2016). Reach tracking reveals dissociable processes underlying cognitive control. *Cognition*, *152*, 114–126. doi:10.1016/j.cognition.2016.03.015

- Erb, C. D., Moher, J., Song, J.-H., & Sobel, D. M. (2017). Cognitive control in action: Tracking the dynamics of rule switching in 5- to 8-year-olds and adults. *Cognition*, *164*, 163–173. doi:10.1016/j.cognition.2017.03.023
- Erb, C. D., Moher, J., Song, J., & Sobel, D. M. (2018). Reach tracking reveals dissociable processes underlying inhibitory control in 5- to 10-year-olds and adults. *Developmental Science*, *21*(2), 1–14. doi:10.1111/desc.12523
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*(3), 215–251. doi:10.1037/0033-295X.87.3.215
- Eysenck, H. J. (1939). Primary mental abilities. *British Journal of Educational Psychology*, *9*, 270-275.
- Fagard, J., & Dahmen, R. (2003). The effects of reading-writing direction on the asymmetry of space perception and directional tendencies: A comparison between French and Tunisian children. *Laterality: Asymmetries of Body, Brain and Cognition*, *8*(1), 39–52. doi:10.1080/713754473
- Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*, *14*(3), 340–347. doi:10.1162/089892902317361886
- Ferrand, L., Ducrot, S., Chausse, P., Maïonchi, P. N., O'Connor, R. J., Parris, B. A., Perret, P., Riggs, K. J., & Augustinova, M. (2019). Stroop interference is a composite phenomenon: Evidence from distinct developmental trajectories of its components. *Developmental Science*. doi:10.1111/desc.12899
- Fishkin, A. S., Kampsnider, J. J., & Pack, L. (1996). Exploring the WISC-III as a measure of giftedness. *Roepers Review: A Journal on Gifted Education*, *18*(3), 226–231. doi:10.1080/02783199609553744
- Fletcher, J. M., Marks, A. G., & Hine, D. W. (2011). Working memory capacity and cognitive styles in decision-making. *Personality And Individual Differences*, *50*(7), 1136-1141. doi:10.1016/j.paid.2011.02.002
- Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, *95*(1), 29–51. doi:10.1037/0033-2909.95.1.29
- Flynn, J. R. (2007). *What is intelligence? Beyond the Flynn effect*. Cambridge University Press.
- Folstein, M. F., Folstein, S. E., McHugh, P. R. (1975). "Mini-mental status". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189-198. doi:10.1016/0022-3956(75)90026-6
- Ford, M. (1995). Two modes of mental representation and problem solution in syllogistic reasoning. *Cognition*, *54*(1), 1–71. doi:10.1016/0010-0277(94)00625-U

- Fournet, N., Roulin, J.-L., Vallet, F., Beaudoin M., Agrigoroaei S., Paignon, A. Dantzer, C., Desrichard O. (2012). Evaluating short-term and working memory in older adults : French normative data. *Aging and Mental Health*, 16(7), 922-30. doi:10.1080/13607863.2012.674487
- Fried, E. I. (2017). The 52 symptoms of major depression: Lack of content overlap among seven common depression scales. *Journal of Affective Disorders*, 208, 191–197. doi:10.1016/j.jad.2016.10.019
- Friedman, D., & Johnson, R., Jr. (2000). Event-related potential (ERP) studies of memory encoding and retrieval: a selective review. *Microscopy Research and Technique*, 51(1), 6–28.
- Friedman, N. P., & Miyake, A. (2004). The reading span test and its predictive power for reading comprehension ability. *Journal Of Memory And Language*, 51(1), 136-158. doi:10.1016/j.jml.2004.03.008
- Fry, A. F., & Hale, S. (2000). Relationships among processing speed, working memory and fluid intelligence in children. *Biological Psychology*, 54(1-3), 1-34. doi:10.1016/s0301-0511(00)00051-x
- Gaillard, V., Barrouillet, P., Jarrold, C., & Camos, V. (2011). Developmental differences in working memory: Where do they come from? *Journal of Experimental Child Psychology*, 110(3), 469–479. doi:10.1016/j.jecp.2011.05.004
- Gathercole, S. E., & Baddeley, A. D. (1993). *Working Memory and Language*. Hove, UK: Erlbaum.
- Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology*, 18(1), 1-16. doi:10.1002/acp.934
- Gaultney, J. F., Bjorklund, D. F., & Goldstein, D. (1996). To Be Young, Gifted, and Strategic: Advantages for Memory Performance. *Journal of Experimental Child Psychology*, 61(1), 43–66. <http://doi.org/10.1006/jecp.1996.0002>
- Geraerts, E., Merckelbach, H., Jelicic, M., & Habets, P. (2007). Suppression of intrusive thoughts and working memory capacity in repressive coping. *The American Journal Of Psychology*, 120(2), 205-218.
- Gevins, A., Smith, M. E., McEvoy, L., & Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. *Cerebral Cortex*, 7(4), 374–385.

- Gimmig, D., Huguet, P., Caverni, J.-P., & Cury, F. (2006). Choking under pressure and working memory capacity: When performance pressure reduces fluid intelligence. *Psychonomic Bulletin & Review*, *13*(6), 1005-1010. doi:10.3758/BF03213916
- Glymour, C. (1998). What went wrong? Reflections on science by observation and the bell curve. *Philosophy of Science*, *65*(1), 1-32.
- Goldhammer, F., Naumann, J., & Greiff, S. (2015). More is not always better: The relation between item response and item response time in Raven's matrices. *Journal of Intelligence*, *3*, 21-40. doi:10.3390/jintelligence3010021
- Gonthier, C. (2020). Charting the diversity of strategic processes in visuo-spatial short-term memory. *Perspectives on Psychological Science*.
- Gonthier, C., Ambrosi, S., & Blaye, A. (en révision). Learning-based before intentional cognitive control: Developmental evidence for a dissociation between implicit and explicit control.
- Gonthier, C., Aubry, A., & Bourdin, B. (2017). Measuring working memory capacity in children using adaptive tasks: Example validation of an adaptive complex span. *Behavior Research Methods*. doi:10.3758/s13428-017-0916-4
- Gonthier, C., Braver, T. S., & Bugg, J. M. (2016). Dissociating proactive and reactive control in the Stroop task. *Memory and Cognition*, *44*(5), 778-788. doi:10.3758/s13421-016-0591-1
- Gonthier, C., & Gavornikova-Baligand, Z. (en révision). Are individual differences in working memory capacity grounded in the reactivation of memory traces ?
- Gonthier, C., Grégoire, J., & Besançon, M. (en révision). No negative Flynn effect in France: the perils of comparing different versions of the same test.
- Gonthier, C., & Hot, P. (2013). Apports de l'électro-encéphalographie à la compréhension de la mémoire. *Revue de neuropsychologie*, *5*(4), 243-254. doi:10.1684/nrp.2013.0280
- Gonthier, C., Macnamara, B., Chow, M., Conway, A. R. A., & Braver, T. S. (2016). Inducing proactive and reactive control shifts in the AX-CPT. *Frontiers in Psychology*, *7*(1822). doi:10.3389/fpsyg.2016.01822
- Gonthier, C., & Roulin, J.-L. (2019). Intraindividual strategy shifts in Raven's matrices, and their dependence on working memory capacity and need for cognition. *Journal of Experimental Psychology: General*. doi:10.1037/xge0000660
- Gonthier, C., & Thomassin, N. (2015). Strategy use fully mediates the relationship between working memory capacity and Raven's matrices. *Journal of Experimental Psychology: General*, *144*(5), 916-924. doi: 10.1037/xge0000101

- Gonthier, C., Thomassin, N., Naëgelé, B., Hot, P., & Detante, O. (2010, Mai). *Analyse des déficits cognitifs consécutifs à un accident vasculaire cérébral*. Journée de la Recherche Médicale, Grenoble, France.
- Gonthier, C., Thomassin, N., & Roulin, J.-L. (2016). The Composite Complex Span : French validation of a short working memory task. *Behavior Research Methods*, *48*(1), 233-242. doi:10.3758/s13428-015-0566-3
- Gonthier, C., Zira, M., Colé, P., & Blaye, A. (2019). Evidencing the developmental shift from reactive to proactive control in early childhood, and its relationship to working memory. *Journal of Experimental Child Psychology*, *177*, 1-16. doi:10.1016/j.jecp.2018.07.001
- Goodhew, S. C. (2020). When cognitive control harms rather than helps: Individuals with high working memory capacity are less efficient at infrequent contraction of attentional breadth. *Psychological Research*. doi:10.1007/s00426-020-01344-x
- Gottfredson, L. S. (1997a). Mainstream science on intelligence: An editorial with 52 signatories, history and bibliography. *Intelligence*, *24*(1), 13–23. doi:10.1016/S0160-2896(97)90011-8
- Gottfredson, L. S. (1997b). Why g matters: The complexity of everyday life. *Intelligence*, *24*(1), 79–132. doi:10.1016/S0160-2896(97)90014-3
- Gottfredson, L. S. (2016). A g theorist on why Kovacs and Conway's process overlap theory amplifies, not opposes, g theory. *Psychological Inquiry*, *27*(3), 210–217. doi:10.1080/1047840X.2016.1203232
- Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation and responses. *Journal of Experimental Psychology: General*, *121*(4), 480–506. doi:10.1037/0096-3445.121.4.480
- Greenfield, P. M. (1997). You can't take it with you: Why ability assessments don't cross cultures. *American Psychologist*, *52*(10), 1115–1124. doi:10.1037/0003-066X.52.10.1115
- Grégoire, J. (1993). Intelligence et vieillissement au WAIS-R Une analyse transversale de l'échantillon d'étalonnage français avec contrôle du niveau scolaire [Intelligence and aging: A cross-sectional analysis of the French standardization sample of the WAIS-R with educational level controlled]. *L'Année Psychologique*, *93*(3), 379–400. doi:10.3406/psy.1993.28701
- Gruber, M. J., & Otten, L. J. (2010). Voluntary control over prestimulus activity related to encoding. *The Journal of Neuroscience*, *30*(29), 9793–9800. doi:10.1523/JNEUROSCI.0915-10.2010

- Guénoilé, F., Speranza, M., Louis, J., Fournernet, P., Revol, O., & Baleyte, J.-M. (2015). Wechsler profiles in referred children with intellectual giftedness: Associations with trait-anxiety, emotional dysregulation, and heterogeneity of Piaget-like reasoning processes. *European Journal of Paediatric Neurology*, *19*, 402-410. doi:10.1016/j.ejpn.2015.03.006
- Hamaker, E. L., Dolan, C. V., & Molenaar, P. C. M. (2005). Statistical Modeling of the Individual: Rationale and Application of Multivariate Stationary Time Series Analysis. *Multivariate Behavioral Research*, *40*(2), 207–233. doi:10.1207/s15327906mbr4002_3
- Hamilton, C. J., Coates, R. O., & Heffernan, T. (2003). What develops in visuo-spatial working memory development? *European Journal of Cognitive Psychology*, *15*(1), 43–69. doi:10.1080/09541440303597
- Hampshire, A., Highfield, R. R., Parkin, B. L., & Owen, A. M. (2012). Fractionating human intelligence. *Neuron*, *76*, 1225-1237. doi:10.1016/j.neuron.2012.06.022
- Hayama, H. R., Johnson, J. D., & Rugg, M. D. (2008). The relationship between the right frontal old/new ERP effect and post-retrieval monitoring: Specific or non-specific? *Neuropsychologia*, *46*(5), 1211–1223. doi:10.1016/j.neuropsychologia.2007.11.021
- Heaton, R. K., Taylor, M. J., & Manly, J. (2003). Demographic effects and use of demographically corrected norms with the WAIS-III and WMS-III. In D. S. Tulsky, D. H. Saklofske, G. J. Chelune, R. K. Heaton, R. J. Ivnik, R. Bornstein, A. Prifitera, & M. F. Ledbetter (Eds.), *Clinical interpretation of the WAIS-III and WMS-III* (pp. 181–210). Academic Press. doi:10.1016/B978-012703570-3/50010-9
- Hedden, T., & Yoon, C. (2006). Individual differences in executive processing predict susceptibility to interference in verbal working memory. *Neuropsychology*, *20*, 511–528. doi:10.1037/0894-4105.20.5.511
- Henderson, D., Poppe, A. B., Barch, D. M., Carter, C. S., Gold, J. M., Ragland, J. D., & ... MacDonald, A. (2012). Optimization of a goal maintenance task for use in clinical applications. *Schizophrenia Bulletin*, *38*(1), 104-113. doi:10.1093/schbul/sbr172
- Henson, R. (2001). Neural working memory. In J. Andrade (Ed.), *Working memory in perspective*. (pp. 151–173). Psychology Press.
- Hill, B. D., Foster, J. D., Elliott, E. M., Shelton, J. T., McCain, J., & Gouvier, W. D. (2013). Need for cognition is related to higher general intelligence, fluid intelligence, and crystallized intelligence, but not working memory. *Journal Of Research In Personality*, *47*(1), 22-25. doi:10.1016/j.jrp.2012.11.001
- Hill, B. D., Foster, J. D., Sofko, C., Elliott, E. M., & Shelton, J. T. (2016). The interaction of ability and motivation: Average working memory is required for Need for Cognition to

- positively benefit intelligence and the effect increases with ability. *Personality And Individual Differences*, 98, 225-228. doi:10.1016/j.paid.2016.04.043
- Hinson, J. M., Jameson, T. L., & Whitney, P. (2002). Somatic markers, working memory, and decision making. *Cognitive, Affective & Behavioral Neuroscience*, 2(4), 341–353. doi:10.3758/CABN.2.4.341
- Hitch, G. J., Halliday, M. S., Dodd, A., & Littler, J. E. (1989). Development of rehearsal in short-term memory: Differences between pictorial and spoken stimuli. *British Journal of Developmental Psychology*, 7(4), 347–362. doi:10.1111/j.2044-835X.1989.tb00811.x
- Holdnack, J. A., Drozdick, L. W., Weiss, L. G., & Iverson, G. L. (2013). *WAIS-IV, WMS-IV, and ACS: Advanced clinical interpretation*. Elsevier Academic Press.
- Hollinger, C. L., & Kosek, S. (1986). Beyond the use of full scale IQ scores. *Gifted Child Quarterly*, 30(2), 74–77. doi:10.1177/001698628603000206
- Hommel, B. (2007). Consciousness and Control: Not Identical Twins. *Journal of Consciousness Studies*, 14(1–2), 155–176.
- Hommel, B. (2013). Dancing in the dark: No role for consciousness in action control. *Frontiers in Psychology*, 4. doi:10.3389/fpsyg.2013.00380
- Horn, J. (1987). A context for understanding information processing studies of human abilities. In P. A. Vernon (Ed.), *Speed of information-processing and intelligence* (pp. 201–238). Ablex Publishing.
- Huguet, P., Brunot, S., & Monteil, J. M. (2001). Geometry versus Drawing: Changing the Meaning of the Task as a Means to Change Performance. *Social Psychology of Education*, 4, 219-234. doi:10.1023/A:1011374700020
- Huguet, P., & Régner, I. (2007). Stereotype threat among schoolgirls in quasi-ordinary classroom circumstances. *Journal of Educational Psychology*, 99(3), 545–560. doi:10.1037/0022-0663.99.3.545
- Hutchison, K. A. (2011). The interactive effects of listwide control, item-based control, and working memory capacity on Stroop performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(4), 851 – 860. doi:10.1037/a0023437
- Hwang, G., Jacobs, J., Geller, A., Danker, J., Sekuler, R., & Kahana, M. J. (2005). EEG correlates of verbal and nonverbal working memory. *Behavioral and Brain Functions*, 1. doi:10.1186/1744-9081-1-20
- Jaeger, A., & Parente, M. A. M. P. (2008). Event-related potentials and the study of memory retrieval: A critical review. *Dementia & Neuropsychologia*, 2(4), 248–255. doi:10.1590/S1980-57642009DN20400003

- Jastrzębski, J., Ciechanowska, I., & Chuderski, A. (2018). The strong link between fluid intelligence and working memory cannot be explained away by strategy use. *Intelligence, 66*, 44–53. doi:10.1016/j.intell.2017.11.002
- Jeffries, S., & Everatt, J. (2004). Working memory: Its role in dyslexia and other specific learning difficulties. *Dyslexia: An International Journal of Research and Practice, 10*(3), 196–214. doi:10.1002/dys.278
- Jensen, A. R. (1987). The g beyond factor analysis. In R. R. Ronning, J. A. Glover, J. C. Conoley, & J. C. Witt (Eds.), *The Influence of Cognitive Psychology on Testing* (pp. 87-142). Lawrence Erlbaum Associates.
- Jensen, A. (1998). *The g factor: The science of mental ability*. Praeger Publishers / Greenwood Publishing Group.
- Johnson, R., Jr., Kreiter, K., Zhu, J., & Russo, B. (1998). A spatio-temporal comparison of semantic and episodic cued recall and recognition using event-related brain potentials. *Cognitive Brain Research, 7*(2), 119–136. doi:10.1016/S0926-6410(98)00017-2
- Johnson, S. B., Blum, R. W., & Giedd, J. N. (2009). Adolescent maturity and the brain: The promise and pitfalls of neuroscience research in adolescent health policy. *Journal of Adolescent Health, 45*(3), 216–221. doi:10.1016/j.jadohealth.2009.05.016
- Jones, D., Beaman, C. P., Macken, W. J. (1996). The object-oriented episodic record model. In S. Gathercole (Ed.), *Models of short-term memory* (pp. 209-238). London: Erlbaum.
- Jones, J. A. H., Sponheim, S. R., & MacDonald, A. W., III. (2010). The dot pattern expectancy task: Reliability and replication of deficits in schizophrenia. *Psychological Assessment, 22*(1), 131–141. doi:10.1037/a0017828
- Jonides, J., Reuter-Lorenz, P. A., Smith, E. E, Awh, E., Barnes, L. L., Drain, M., Glass, J., Lauber, E. J., Patalano, A. L., & Schumacher, E. H. (1996). Verbal and spatial working memory in humans. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory, Vol. 35* (pp. 43-88). San Diego, CA, US: Academic Press.
- Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 29-61). London, UK: Academic Press.
- Kan, K.-J., van der Maas, H. L. J., & Kievit, R. A. (2016). Process overlap theory: Strengths, limitations, and challenges. *Psychological Inquiry, 27*(3), 220–228. doi:10.1080/1047840X.2016.1182000
- Kane, M. J., Conway, A. R. A., Miura, T. K., & Colflesh, G. J. H. (2007). Working memory, attention control, and the n-back task: A question of construct validity. *Journal of*

- Experimental Psychology: Learning, Memory, and Cognition*, 33(3), 615–622. doi:10.1037/0278-7393.33.3.615
- Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(2), 336-358. doi: 10.10371/0278-7393.26.2.336
- Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, 132(1), 47-70. doi:10.1037/0096-3445.132.1.47
- Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, 133(2), 189-217. doi:10.1037/0096-3445.133.2.189
- Kaufman, A. S. (2010). "In what way are apples and oranges alike?" A critique of Flynn's interpretation of the Flynn effect. *Journal of Psychoeducational Assessment*, 28(5), 382–398. doi:10.1177/0734282910373346
- Kidd, A. H. (1962). The Culture-Fair Aspects of Cattell's Test of g: Culture-Free. *The Journal of Genetic Psychology*, 101(2), 343–362. doi:10.1080/00221325.1962.10533635
- Klauer, K., & Zhao, Z. (2004). Double Dissociations in Visual and Spatial Short-Term Memory. *Journal Of Experimental Psychology: General*, 133(3), 355-381. doi:10.1037/0096-3445.133.3.355
- Klayman, J. (1985). Children's decision strategies and their adaptation to task characteristics. *Organizational Behavior And Human Decision Processes*, 35(2), 179-201. doi:10.1016/0749-5978(85)90034-2
- Klopfer, D. S. (1996). Stroop interference and color-word similarity. *Psychological Science*, 7(3), 150–157. doi:10.1111/j.1467-9280.1996.tb00348.x
- Kovacs, K., & Conway, A. R. A. (2016). Process overlap theory: A unified account of the general factor of intelligence. *Psychological Inquiry*, 27(3), 151–177. doi:10.1080/1047840X.2016.1153946
- Kranzler, J. H., & Jensen, A. R. (1991). The nature of psychometric g: Unitary process or a number of independent processes? *Intelligence*, 15(4), 397–422. doi:10.1016/0160-2896(91)90003-V

- Kroll, N. E., Parks, T., Parkinson, S. R., Bieber, S. L., & Johnson, A. (1970). Short-term memory while shadowing: Recall of visually and of aurally presented letters. *Journal Of Experimental Psychology*, *85*(2), 220-224. doi:10.1037/h0029544
- Krumm, S., Schmidt-Atzert, L., Buehner, M., Ziegler, M., Michalczyk, K., & Arrow, K. (2009). Storage and non-storage components of working memory predicting reasoning: A simultaneous examination of a wide range of ability factors. *Intelligence*, *37*(4), 347–364. doi:10.1016/j.intell.2009.02.003
- Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2004). Academic Performance, Career Potential, Creativity, and Job Performance: Can One Construct Predict Them All? *Journal of Personality and Social Psychology*, *86*(1), 148–161. doi:10.1037/0022-3514.86.1.148
- Kyllonen, P. C. (2002). g: Knowledge, speed, strategies, or working-memory capacity? A systems perspective. In R. J. Sternberg & E. L. Grigorenko (Eds.), *The general factor of intelligence: How general is it?* (p. 415–445). Lawrence Erlbaum Associates Publishers.
- Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, *14*(4), 389-433. doi:10.1016/s0160-2896(05)80012-1
- Labouret, G., & Grégoire, J. (2018). La dispersion intra-individuelle et le profil des scores dans les QI élevés. *A.N.A.E.*, *154*, 271-279.
- Lamiell, J. T. (1981). Toward an idiothetic psychology of personality. *American Psychologist*, *36*(3), 276–289. doi:10.1037/0003-066X.36.3.276
- Larson, G. E., & Saccuzzo, D. P. (1989). Cognitive correlates of general intelligence: Toward a process theory of g. *Intelligence*, *13*(1), 5–31. doi:10.1016/0160-2896(89)90003-2
- Lautrey, J. (2003). La psychologie différentielle à l'épreuve de la variabilité intraindividuelle. In A. Vom Hofe, H. Charvin, J.-L. Bernaud, & D. Guédon (Eds.), *Psychologie différentielle – recherches et réflexions* (pp. 9-28). Presses Universitaires de Rennes.
- Lehle, C., & Hübner, R. (2008). On-the-fly adaptation of selectivity in the flanker task. *Psychonomic Bulletin & Review*, *15*(4), 814–818. doi:10.3758/PBR.15.4.814
- Lehrer, R., Guckenberger, T., & Lee, O. (1988). Comparative study of the cognitive consequences of inquiry-based Logo instruction. *Journal Of Educational Psychology*, *80*(4), 543-553. doi:10.1037/0022-0663.80.4.543
- Lemaire, P., & Reder, L. (1999). What affects strategy selection in arithmetic? The example of parity and five effects on product verification. *Memory & Cognition*, *27*(2), 364–382. doi:10.3758/BF03211420
- Lemaire, P., & Siegler, R. S. (1995). Four aspects of strategic change: Contributions to children's learning of multiplication. *Journal of Experimental Psychology: General*, *124*(1), 83-97. doi:10.1037/0096-3445.124.1.83

- Lépine, R., Barrouillet, P., & Camos, V. (2005). What makes working memory spans so predictive of high-level cognition? *Psychonomic Bulletin & Review*, *12*(1), 165 – 170. doi:10.3758/BF03196363
- Lezak, M. D. (1982). The problem of assessing executive functions. *International Journal of Psychology*, *17*(2-3), 281-297. doi:10.1080/00207598208247445
- Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2013). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review*, doi:10.3758/s13423-013-0565-2
- Liu, T., Xiao, T., Shi, J., & Zhao, D. (2011). Response preparation and cognitive control of highly intelligent children: a Go-Nogo event-related potential study. *Neuroscience*, *180*(C), 122–128. <http://doi.org/j.neuroscience.2011.02.022>
- Loesche, P., Wiley, J., & Hasselhorn, M. (2015). How knowing the rules affects solving the Raven Advanced Progressive Matrices Test. *Intelligence*, *48*, 58–75. doi:10.1016/j.intell.2014.10.004
- Logan, G. D., & Zbrodoff, N. (1979). When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Memory & Cognition*, *7*(3), 166 – 174. doi:10.3758/BF03197535
- Logie, R. H. (1995). *Visuo-spatial working memory*. Lawrence Erlbaum Associates, Inc.
- Logie, R. H. (2016). Retiring the central executive. *The Quarterly Journal of Experimental Psychology*, *69*(10), 2093–2109. doi:10.1080/17470218.2015.1136657
- Logie, R. H. (2018). Human cognition: Common principles and individual variation. *Journal of Applied Research in Memory and Cognition*, *7*(4), 471–486. doi:10.1016/j.jarmac.2018.08.001
- Loosli, S. V., Rahm, B., Unterrainer, J. M., Weiller, C., & Kaller, C. P. (2014). Developmental change in proactive interference across the life span: Evidence from two working memory tasks. *Developmental Psychology*, *50*(4), 1060 – 1072. doi:10.1037/a0035231
- Lord, F. M. (1968). Some test theory for tailored testing. *ETS Research Bulletin Series*, *1968*(2), i – 62. doi:10.1002/j.2333-8504.1968.tb00562.x
- Lucidi, A., Loaiza, V., Camos, V., & Barrouillet, P. (2014). Assessing working memory capacity through time-constrained elementary activities. *Journal of General Psychology*, *141*(2), 98–112. doi:10.1080/00221309.2013.870121
- Luria, A. R., & Tsvetkova, L. S. (1964). The programming of constructive activity in local brain injuries. *Neuropsychologia*, *2*(2), 95-107. doi: 10.1016/0028-3932(64)90015-6

- Lustig, C., May, C. P., & Hasher, L. (2001). Working memory span and the role of proactive interference. *Journal Of Experimental Psychology: General*, *130*(2), 199 – 207. doi:10.1037/0096-344
- Lynn, R. (2015). *Race differences in intelligence: An evolutionary analysis*. Washington Summit Publishers.
- Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, *17*(3), 347–356. doi:10.1038/nn.3655
- MacDonald, A. W., III, Goghari, V. M., Hicks, B. M., Flory, J. D., Carter, C. S., & Manuck, S. B. (2005). A convergent-divergent approach to context processing, general intellectual functioning, and the genetic liability to schizophrenia. *Neuropsychology*, *19*(6), 814–821. doi:10.1037/0894-4105.19.6.814
- MacLeod, C. M., Dodd, M. D., Sheard, E. D., Wilson, D. E., & Bibi, U. (2003). In opposition to inhibition. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory, Vol. 43* (pp. 163-214). New York, NY, US: Elsevier Science.
- Macmillan, M. (2008). Phineas Gage - Unravelling the myth. *The Psychologist*, *21*(9), 828-831.
- Majerus, S. (2013). Language repetition and short-term memory: An integrative framework. *Frontiers in Human Neuroscience*, *7*. doi:10.3389/fnhum.2013.00357
- Majerus, S., D'Argembeau, A., Perez, T. M., Belayachi, S., Van der Linden, M., Collette, F., Salmon, E., Seurinck, R., Fias, W., & Maquet, P. (2010). The commonality of neural networks for verbal and visual short-term memory. *Journal of Cognitive Neuroscience*, *22*(11), 2570–2593. doi:10.1162/jocn.2009.21378
- Mäki-Marttunen, V., Hagen, T., & Espeseth, T. (2019). Proactive and reactive modes of cognitive control can operate independently and simultaneously. *Acta Psychologica*, *199*. doi:10.1016/j.actpsy.2019.102891
- Manes, F., Sahakian, B., Clark, L., Rogers, R., Antoun, N., Aitken, M., & Robbins, T. (2002). Decision-making processes following damage to the prefrontal cortex. *Brain: A Journal of Neurology*, *125*(3), 624–639. doi:10.1093/brain/awf049
- Marr, D. (1976). Early processing of visual information. *Philosophical Transactions of the Royal Society of London: Series B*, *275*, 483-524.
- Martinussen, R., Hayden, J., Hogg-Johnson, S., & Tannock, R. (2005). A Meta-Analysis of Working Memory Impairments in Children With Attention-Deficit/Hyperactivity Disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *44*(4), 377–384. doi:10.1097/01.chi.0000153228.72591.73

- McEvoy, L. K., Smith, M. E., & Gervins, A. (1998). Dynamic cortical networks of verbal and spatial working memory: Effects of memory load and task practice. *Cerebral Cortex*, *8*(7), 563–574. doi:10.1093/cercor/8.7.563
- McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*(1), 1-10. doi:10.1016/j.intell.2008.08.004
- McNamara, D. S., & Scott, J. L. (2001). Working memory capacity and strategy use. *Memory & Cognition*, *29*(1), 10-17. doi:10.3758/BF03195736
- Mecklinger, A., & Müller, N. N. (1996). Dissociations in the processing of 'what' and 'where' information in working memory: An event-related potential analysis. *Journal Of Cognitive Neuroscience*, *8*(5), 453-473. doi:10.1162/jocn.1996.8.5.453
- Meijer, J., & Oostdam, R. (2007). Test anxiety and intelligence testing: A closer examination of the stage-fright hypothesis and the influence of stressful instruction. *Anxiety, Stress & Coping: An International Journal*, *20*(1), 77–91. doi:10.1080/10615800600967862
- Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, *49*, 270–291. doi:10.1037/a0028228
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review Of Neuroscience*, *24*, 167-202. doi:10.1146/annurev.neuro.24.1.167
- Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York, NY, US: Henry Holt and Co. doi:10.1037/10039-000
- Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, *102*(2), 246–268. doi:10.1037/0033-295X.102.2.246
- Mitchum, A. L., & Kelley, C. M. (2010). Solve the problem first: Constructive solution strategies can influence the accuracy of retrospective confidence judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(3), 699. [https://doi.org/ 10.1037/a0019182](https://doi.org/10.1037/a0019182)
- Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. doi:10.1177/0963721411429458
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex 'frontal lobe' tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49-100. doi:10.1006/cogp.1999.0734

- Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement: Interdisciplinary Research and Perspectives*, 2(4), 201–218. doi:10.1207/s15366359mea0204_1
- Mora, G., & Camos, V. (2013). Two systems of maintenance in verbal working memory: Evidence from the word length effect. *Plos ONE*, 8(7):e70026. doi:10.1371/journal.pone.0070026
- Morey, C. C. (2018). The case against specialized visual-spatial short-term memory. *Psychological Bulletin*, 144(8), 849–883. doi:10.1037/bul0000155
- Msetfi, R. M., Murphy, R. A., Kornbrot, D. E., & Simpson, J. (2009). Impaired context maintenance in mild to moderately depressed students. *The Quarterly Journal of Experimental Psychology*, 62(4), 653–662. doi:10.1080/17470210802486092
- Mulholland, T. M., Pellegrino, J. W., & Glaser, R. (1980). Components of geometric analogy solution. *Cognitive Psychology*, 12(2), 252–284. doi:10.1016/0010-0285(80)90011-0
- Munakata, Y., Snyder, H. R., & Chatham, C. H. (2012). Developing cognitive control: Three key transitions. *Current Directions In Psychological Science*, 21(2), 71–77. doi:10.1177/0963721412436807
- Naëgelé, B., & Mazza, S. (2003). *Le PASAT modifié*. Editions Solal.
- Newell, A., & Simon, H. A. (1956). *The logic theory machine: A complex information processing system*. Rand Corporation.
- Noreen, S., Cooke, R., & Ridout, N. (2019). Investigating the mediating effect of working memory on intentional forgetting in dysphoria. *Psychological Research*. doi:10.1007/s00426-019-01225-y
- Norman, D. A., & Shallice, T. (1986). Attention to Action. In R. Davidson, G. Schwartz, & D. Shapiro (Eds.), *Consciousness and Self-Regulation* (Vol. 4, pp. 1–18). New York, NY, US: Plenum.
- Nunes Carraher, T., Carraher, D. W., & Schliemann, A. D. (1985). Mathematics in the streets and in schools. *British Journal of Developmental Psychology*, 3, 21–29.
- Oberauer, K., Schulze, R., Wilhelm, O., & Süß, H.-M. (2005). Working Memory and Intelligence – Their Correlation and Their Relation: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, 131(1), 61–65. doi:10.1037/0033-2909.131.1.61
- Oberauer, K., Süß, H.-M., Wilhelm, O., & Wittman, W. W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence*, 31(2), 167–193. doi:10.1016/S0160-2896(02)00115-0

- Oostdam, R., & Meijer, J. (2003). Influence of test anxiety on measurement of intelligence. *Psychological Reports, 92*(1), 3–20. doi:10.2466/PRO.92.1.3-20
- Otten, L. J., Quayle, A. H., Akram, S., Ditewig, T. A., & Rugg, M. D. (2006). Brain activity before an event predicts later recollection. *Nature Neuroscience, 9*(4), 489–491. doi:10.1038/nn1663
- Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-Back Working Memory Paradigm: A Meta-Analysis of Normative Functional Neuroimaging Studies. *Human Brain Mapping, 25*(1), 46-59. doi:10.1002/hbm.20131
- Partchev, I., & De Boeck, P. (2012). Can fast and slow intelligence be differentiated? *Intelligence, 40*, 23–32. doi:10.1016/j.intell.2011.11.002
- Paulesu, E., Frith, C. D., & Frackowiak, R. S. (1993). The neural correlates of the verbal component of working memory. *Nature, 362*(6418), 342–345. doi:10.1038/362342a0
- Paxton, J. L., Barch, D. M., Racine, C. A., & Braver, T. S. (2008). Cognitive control, goal maintenance, and prefrontal function in healthy aging. *Cerebral Cortex, 18*(5), 1010-1028. doi:10.1093/cercor/bhm135
- Paxton, J. L., Barch, D. M., Storandt, M., & Braver, T. S. (2006). Effects of environmental support and strategy training on older adults' use of context. *Psychology and Aging, 21*(3), 499-509. doi:10.1037/0882-7974.21.3.499
- Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance, 16*, 366–387. doi:10.1016/0030-5073(76)90022-2
- Pearce, M. S., Deary, I. J., Young, A. H., & Parker, L. (2005). Growth in early life and childhood IQ at age 11 years: the Newcastle Thousand Families Study. *International Journal of Epidemiology, 34*, 673–677. doi:10.1093/ije/dyi038
- Perret, P., & Dauvier, B. (2018). Children's allocation of study time during the solution of Raven's progressive matrices. *Journal of Intelligence, 6*(1). doi:10.3390/jintelligence6010000
- Perret, P., & Dauvier, B. (2020). "Trop face" les matrices de Raven ? Le rôle de la clairvoyance métacognitive dans la régulation du raisonnement chez l'enfant [Chapitre en préparation].
- Petrides, M., & Milner, B. (1982). Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia, 20*(3), 249–262. doi:10.1016/0028-3932(82)90100-2

- Pietschnig, J., & Voracek, M. (2015). One century of global IQ gains: A formal meta-analysis of the Flynn effect (1909–2013). *Perspectives on Psychological Science, 10*(3), 282–306. doi:10.1177/1745691615577701
- Pontifex, M. B., Hillman, C. H., Fernhall, B., Thompson, K. M., & Valentini, T. A. (2009). The effect of acute aerobic and resistance exercise on working memory. *Medicine & Science in Sports & Exercise, 41*(4), 927–934. doi:10.1249/MSS.0b013e3181907d69
- Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience, 139*(1), 23–38. doi:10.1016/j.neuroscience.2005.06.005
- Postle, B. R., Idzikowski, C., Della Sala, S., Logie, R. H., & Baddeley, A. D. (2006). The selective disruption of spatial working memory by eye movements. *The Quarterly Journal Of Experimental Psychology, 59*(1), 100–120. doi:10.1080/17470210500151410
- Rabbitt, P. (1997). Methodologies and models in the study of executive function. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 1–38). East Sussex, UK: Psychology Press Publishers.
- Raven, J. (2000). The Raven's Progressive Matrices: Change and stability over culture and time. *Cognitive Psychology, 41*(1), 1–48. doi:10.1006/cogp.1999.0735
- Raven, J. (2008). The Raven Progressive Matrices tests: Their theoretical basis and measurement model. In J. Raven & J. C. Raven (Eds.), *Uses and Abuses of Intelligence: Studies Advancing Spearman and Raven's Quest for Non-Arbitrary Metrics*. Royal Fireworks Press.
- Reder, L. M., & Schunn, C. D. (1999). Bringing together the psychometric and strategy worlds: Predicting adaptivity in a dynamic task. In D. Gopher, A. Koriat, D. Gopher, A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 315–342). The MIT Press.
- Redick, T. S. (2014). Cognitive control in context: Working memory capacity and proactive control. *Acta Psychologica, 145*, 1–9. doi:10.1016/j.actpsy.2013.10.010
- Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal Of Psychological Assessment, 28*(3), 164–171. doi:10.1027/1015-5759/a000123
- Redick, T. S., & Engle, R. W. (2011). Integrating working memory capacity and context-processing views of cognitive control. *The Quarterly Journal of Experimental Psychology, 64*(6), 1048–1055. doi:10.1080/17470218.2011.577226

- Redick, T. S., & Lindsey, D. R. B. (2013). Complex span and n-back measures of working memory: A meta-analysis. *Psychonomic Bulletin & Review*, 20(6), 1102–1113. doi:10.3758/s13423-013-0453-9
- Reimers, S., & Maylor, E. A. (2005). Task Switching Across the Life Span: Effects of Age on General and Specific Switch Costs. *Developmental Psychology*, 41(4), 661–671. doi:10.1037/0012-1649.41.4.661
- Reuchlin, M. (1978). Processus vicariants et différences individuelles. [Vicariant processes and individual differences]. *Journal de Psychologie Normale et Pathologique*, 75(2), 133-145.
- Rhodes, S., Abbene, E. E., Meierhofer, A. M., & Naveh-Benjamin, M. (2020). Age differences in the precision of memory at short and long delays. *Psychology and Aging*. doi:10.1037/pag0000565
- Richmond, L., Redick, T. S., & Braver, T. S. (2013). Remembering to prepare: The benefits (and costs) of high working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(6), 1764-1777. doi:10.1037/xlm0000122
- Rivollier, G., Quinton, J.-C., Gonthier, C., & Smeding, A. (2020). Looking with the computer mouse: How to unveil problem-solving strategies in matrix reasoning without eye-tracking. *Behavior Research Methods*.
- Roberts, B. M., Hsieh, L.-T., & Ranganath, C. (2013). Oscillatory activity during maintenance of spatial and temporal information in working memory. *Neuropsychologia*, 51(2), 349–357. doi:10.1016/j.neuropsychologia.2012.10.009
- Robison, M. K., & Brewer, G. A. (2020). Individual differences in working memory capacity and the regulation of arousal. *Attention, Perception, & Psychophysics*. doi:10.3758/s13414-020-02077-0
- Rodríguez-Naveiras, E., Verche, E., Hernández-Lastiri, P., Montero, R., & Borges, Á. (2019). Differences in working memory between gifted or talented students and community samples: A meta-analysis. *Psicothema*, 31(3), 255-262. doi:10.7334/psicothema2019.18
- Rose, N. S. (2013). Individual differences in working memory, secondary memory, and fluid intelligence: Evidence from the levels-of-processing span task. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, 67(4), 260–270. doi:10.1037/a0034351
- Rose, N. S., & Craik, F. I. M. (2012). A processing approach to the working memory/long-term memory distinction: Evidence from the levels-of-processing span task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(4), 1019–1029. doi:10.1037/a0026976

- Rosvold, H., Mirsky, A. F., Sarason, I., Bransome, E. R., and Beck, L. H. (1956). A continuous performance test of brain damage. *Journal of Consulting Psychology, 20*(5), 343–350. doi: 10.1037/h0043220
- Rouder, J. N., Morey, R. D., Morey, C. C., & Cowan, N. (2011). How to measure working memory capacity in the change detection paradigm. *Psychonomic Bulletin & Review, 18*(2), 324–330. doi:10.3758/s13423-011-0055-3
- Rozencwajg, P. (1991). Analysis of problem solving strategies on the Kohs block design test. *European Journal of Psychology of Education, 6*(1), 73-88. doi:10.1007/BF03173149
- Rozencwajg, P., Aliamer, V., & Ombredane, E. (2009). Le fonctionnement cognitif d'enfants atypiques à travers leur QI [Cognitive functioning in atypical children in regard to their IQ]. *Pratiques Psychologiques, 15*, 343-365. doi:10.1016/j.prps.2008.02.003
- Rozencwajg, P., Cherfi, M., Ferrandez, A. M., Lautrey, J., Lemoine, C., & Loarer, E. (2005). Age-related differences in the strategies used by middle-aged adults to solve a block design task. *The International Journal of Aging and Human Development, 60*(2), 159-182. doi:10.2190/HOAR-68HR-RRPE-LRBH
- Rozencwajg, P., & Corroyer, D. (2002). Strategy development in a block design task. *Intelligence, 30*(1), 1–25. doi:10.1016/S0160-2896(01)00063-0
- Ruchkin, D. S., Berndt, R. S., Johnson, R., Jr., Ritter, W., Grafman, J., & Canoune, H. L. (1997). Modality-specific processing streams in verbal working memory: Evidence from spatio-temporal patterns of brain activity. *Cognitive Brain Research, 6*(2), 95–113. doi:10.1016/S0926-6410(97)00021-9
- Rugg, M. D., & Wilding, E. L. (2000). Retrieval processing and episodic memory. *Trends in Cognitive Sciences, 4*(3), 108-115. doi:10.1016/s1364-6613(00)01445-5.
- Rundquist, E. A. (1936). Intelligence test scores and school marks of high school seniors in 1929 and 1933. *School & Society, 43*, 301–304.
- Salthouse, T. A. (1993). Influence of working memory on adult age differences in matrix reasoning. *British Journal of Psychology, 84*(2), 171-199. doi:10.1111/j.2044-8295.1993.tb02472.x
- Scharfen, J., Jansen, K., & Holling, H. (2018). Retest effects in working memory capacity tests: A meta-analysis. *Psychonomic Bulletin & Review, 25*(6), 2175–2199. doi:10.3758/s13423-018-1461-6
- Schelble, J. L., Therriault, D. J., & Miller, M. D. (2012). Classifying retrieval strategies as a function of working memory. *Memory & Cognition, 40*(2), 218-230. doi:10.3758/s13421-011-0149-1

- Schmeichel, B. J., & Tang, D. (2014). The relationship between individual differences in executive functioning and emotion regulation: A comprehensive review. In J. P. Forgas & E. Harmon-Jones (Eds.), *Motivation and its regulation: The control within*. (pp. 133–151). Psychology Press.
- Schmidt, J. R., & Besner, D. (2008). The Stroop effect: Why proportion congruent has nothing to do with congruency and everything to do with contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(3), 514–523. doi:10.1037/0278-7393.34.3.514
- Schmiedek, F., Lövdén, M., & Lindenberger, U. (2014). A task is a task is a task: Putting complex span, n-back, and other working memory indicators in psychometric context. *Frontiers in Psychology*, *5*. doi:10.3389/fpsyg.2014.01475
- Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll model of intelligence. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (p. 99–144). The Guilford Press.
- Schunn, C. D., & Reder, L. M. (1998). Strategy adaptivity and individual differences. In D. L. Medin, D. L. Medin (Eds.), *The psychology of learning and motivation: Advances in research and theory, Vol. 38* (pp. 115-154). Academic Press.
- Schunn, C. D., & Reder, L. M. (2001). Another source of individual differences: Strategy adaptivity to changing rates of success. *Journal Of Experimental Psychology: General*, *130*(1), 59-76. doi:10.1037/0096-3445.130.1.59
- Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. Oxford University Press.
- Siegler, R. S. (2016). Continuity and change in the field of cognitive development and in the perspectives of one cognitive developmentalist. *Child Development Perspectives*, *10*(2), 128–133. doi:10.1111/cdep.12173
- Silver, S. J., & Clampit, M. K. (1990). WISC-R profiles of high ability children: Interpretation of verbal-performance discrepancies. *Gifted Child Quarterly*, *34*(2), 76–79. doi:10.1177/001698629003400205
- Simmering, V. R., & Perone, S. (2013). Working memory capacity as a dynamic process. *Frontiers in Psychology*, *3*. doi:10.3389/fpsyg.2012.00567
- Smith, E. E., Jonides, J., & Koeppel, R. A. (1996). Dissociating verbal and spatial working memory using PET. *Cerebral Cortex*, *6*(1), 11-20. doi:10.1093/cercor/6.1.11
- Snow, R. E. (1978). Eye fixation and strategy analyses of individual differences in cognitive aptitudes. In A. M. Lesgold, J. W. Pellegrino, S. D. Fokkema, & R. Glaser (Eds.), *Cognitive psychology and instruction* (pp. 299-308). New York: Plenum Press.

- Snow, R. E. (1980). Aptitude processes. *Aptitude, Learning, and Instruction*, 1, 27–63.
- Snow, R. E. (1992). Aptitude theory: Yesterday, today, and tomorrow. *Educational Psychologist*, 27(1), 5–32. doi:10.1207/s15326985ep2701_3
- Spanoudis, G., & Demetriou, A. (2020). Mapping mind-brain development: Towards a comprehensive theory. *Journal of Intelligence*, 8(19). doi:10.3390/jintelligence8020019
- Spearman, C. (1904). "General intelligence," objectively determined and measured. *The American Journal of Psychology*, 15(2), 201–293. doi:10.2307/1412107
- Spearman, C. (1927). *The abilities of man: Their nature and measurement*. Macmillan.
- St Clair-Thompson, H. L. (2007). The influence of strategies upon relationships between working memory and cognitive skills. *Memory*, 15(4), 353-365. doi:10.1080/09658210701261845
- St Clair-Thompson, H. L. (2010). Backwards digit recall: A measure of short-term memory or working memory? *European Journal of Cognitive Psychology*, 22(2), 286–296. doi:10.1080/09541440902771299
- Sternberg, S. (1969). Memory-Scanning: Mental Processes Revealed by Reaction-Time Experiments. *American Scientist*, 57(4), 421-457.
- Sternberg, R. J. (1999). Intelligence as developing expertise. *Contemporary Educational Psychology*, 24(4), 359–375. doi:10.1006/ceps.1998.0998
- Sternberg, R. J., & Rifkin, B. (1979). The development of analogical reasoning processes. *Journal of Experimental Child Psychology*, 27(2), 195–232. doi:10.1016/0022-0965(79)90044-4
- Strenze, T. (2007). Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. *Intelligence*, 35(5), 401–426. doi:10.1016/j.intell.2006.09.004
- Stuss, D. T. (2011). Functions of the frontal lobes: Relation to executive functions. *Journal of the International Neuropsychological Society*, 17(5), 759–765. doi:10.1017/S1355617711000695
- Süß, H.-M., Oberauer, K., Wittmann, W. W., Wilhelm, O., & Schulze, R. (2002). Working-memory capacity explains reasoning ability: and a little bit more. *Intelligence*, 30(3), 261-288. doi:10.1016/S0160-2896(01)00100-3
- Tafuro, A., Vallesi, A., Gonthier, C., & Ambrosini, E. (en révision). Tracking proactive and reactive cognitive control in interference resolution.
- Thomas, D., & Zumbo, B. D. (2012). Difference scores from the point of view of reliability and repeated-measures ANOVA: In defense of difference scores for data

- analysis. *Educational And Psychological Measurement*, 72(1), 37-43. doi:10.1177/0013164411409929
- Thomassin, N., Gonthier, C., Guerraz, M., & Roulin, J.-L. (2015). The hard fall effect: high working memory capacity leads to a higher, but less robust short-term memory performance. *Experimental Psychology*, 62(2), 89-97. doi:10.1027/1618-3169/a000276
- Thomson, G. H. (1916). A hierarchy without a general factor. *British Journal of Psychology*, 8(3), 271-281. doi:10.1111/j.2044-8295.1916.tb00133.x
- Thurstone, L. L. (1935). *Vectors of mind*. University of Chicago Press.
- Thurstone, L. L. (1940). Current issues in factor analysis. *Psychological Bulletin*, 37(4), 189–236. doi:10.1037/h0059402
- Trahan, L. H., Stuebing, K. K., Fletcher, J. M., & Hiscock, M. (2014). The Flynn effect: A meta-analysis. *Psychological Bulletin*, 140(5), 1332–1360. doi:10.1037/a0037173
- Turley-Ames, K. J., & Whitfield, M. M. (2003). Strategy training and working memory task performance. *Journal of Memory and Language*, 49(4), 446-468. doi:10.1016/s0749-596x(03)00095-0
- Turnbull, O. H., Evans, C. E. Y., Bunce, A., Carzolio, B., & O'Connor, J. (2005). Emotion-based learning and central executive resources: An investigation of intuition and the Iowa Gambling Task. *Brain and Cognition*, 57(3), 244–247. doi:10.1016/j.bandc.2004.08.053
- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, 28(2), 127-154. doi:10.1016/0749-596x(89)90040-5
- Unsworth, N., & Engle, R. W. (2005). Working memory capacity and fluid abilities: Examining the correlation between Operation Span and Raven. *Intelligence*, 33(1), 67-81. doi:10.1016/j.intell.2004.08.003
- Unsworth, N., & Engle, R. W. (2006). A temporal-contextual retrieval account of complex span: An analysis of errors. *Journal of Memory and Language*, 54(3), 346–362. doi:10.1016/j.jml.2005.11.004
- Unsworth, N., & Engle, R. W. (2007a). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, 114(1), 104-132. doi:10.1037/0033-295X.114.1.104
- Unsworth, N., & Engle, R. W. (2007b). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin*, 133(6), 1038-1066. doi:10.1037/0033-2909.133.6.1038

- Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, 37(3), 498-505. doi:10.3758/bf03192720
- Van de Vijver, F. J. R. (2016). Assessment in education in multicultural populations. In G. T. L. Brown & L. R. Harris (Eds.), *Handbook of human and social conditions in assessment* (pp. 436-453). Routledge.
- Van der Linden, M., Poncelet, M., & Majerus, S. (2007). Working memory dysfunctions in stroke patients. In O. Godefroy (Ed.), *The behavioral and cognitive neurology of stroke* (pp. 431-443). Cambridge University Press.
- van der Maas, H. L., Dolan, C. V., Grasman, R. P., Wicherts, J. M., Huizenga, H. M., Raijmakers, M. E. (2006). A dynamical model of general intelligence: the positive manifold of intelligence by mutualism. *Psychological Review*, 113(4), 842-861. doi:10.1037/0033-295X.113.4.842
- van Gaal, S., de Lange, F. P., & Cohen, M. X. (2012). The role of consciousness in cognitive control and decision making. *Frontiers in Human Neuroscience*, 6. doi:10.3389/fnhum.2012.00121
- Vandierendonck, A., De Vooght, G., & Van der Goten, K. (1998). Interfering with the central executive by means of a random interval repetition task. *The Quarterly Journal Of Experimental Psychology A: Human Experimental Psychology*, 51A(1), 197 – 218. doi:10.1080/027249898391828
- Vernon, P. A. (1983). Speed of information processing and general intelligence. *Intelligence*, 7(1), 53-70. doi:10.1016/0160-2896(83)90006-5
- Vigneau, F., Caissie, A. F., & Bors, D. A. (2006). Eye-movement analysis demonstrates strategic influences on intelligence. *Intelligence*, 34(3), 261–272. doi:10.1016/j.intell.2005.11.003
- Vock, M., & Holling, H. (2008). The measurement of visuo-spatial and verbal-numerical working memory: Development of IRT-based scales. *Intelligence*, 36(2), 161 – 182. doi:10.1016/j.intell.2007.02.004
- Wang, P. P., & Bellugi, U. (1994). Evidence from two genetic syndromes for a dissociation between verbal and visual-spatial short-term memory. *Journal Of Clinical And Experimental Neuropsychology*, 16(2), 317-322. doi:10.1080/01688639408402641
- Wechsler, D. (2011). *Manuel de l'Echelle d'Intelligence de Wechsler Pour Adultes - 4ème édition* [Manual for the Wechsler Adult Intelligence Scale - Fourth Edition]. ECPA par Pearson.

- Weiss, D. J. (1974). *Strategies of adaptive ability measurement*. Oxford, England: University of Minnesota, Department of Psychology.
- Wendt, M., & Luna-Rodriguez, A. (2009). Conflict frequency affects flanker interference: Role of stimulus-ensemble-specific practice and flanker-response contingencies. *Experimental Psychology*, *56*(3), 206–217. doi:10.1027/1618-3169.56.3.206
- Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I Latent structure. *Developmental Psychology*, *44*(2), 575–587. doi:10.1037/0012-1649.44.2.575
- Wiley, J., & Jarosz, A. F. (2012). How working memory capacity affects problem solving. In B. H. Ross (Ed.), *The psychology of learning and motivation (Vol 56)* (pp. 185-227). San Diego, CA, US: Elsevier Academic Press. doi:10.1016/B978-0-12-394393-4.00006-6
- Wilhelm, O., Hildebrandt, A., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Frontiers in Psychology*, *4*. doi:10.3389/fpsyg.2013.00433
- Wilkinson, S. C. (1993). WISC—R profiles of children with superior intellectual ability. *Gifted Child Quarterly*, *37*(2), 84–91. doi:10.1177/001698629303700206
- Wilkinson, L., Scholey, A., & Wesnes, K. (2002). Chewing gum selectively improves aspects of memory in healthy volunteers. *Appetite*, *38*(3), 235–236. doi:10.1006/appe.2002.0473
- Wongupparaj, P., Kumari, V., & Morris, R. G. (2015). A cross-temporal meta-analysis of Raven's Progressive Matrices: Age groups and developing versus developed countries. *Intelligence*, *49*, 1–9. doi:10.1016/j.intell.2014.11.008
- Wood, G., Vine, S. J., & Wilson, M. R. (2016). Working memory capacity, controlled attention and aiming performance under pressure. *Psychological Research*, *80*(4), 510–517. doi:10.1007/s00426-015-0673-x
- Wood, J. N. (2011). When do spatial and visual working memory interact? *Attention, Perception, & Psychophysics*, *73*(2), 420–439. doi:10.3758/s13414-010-0048-8
- Woodley of Menie, M. A., & Dunkel, C. S. (2015). In France, are secular IQ losses biologically caused? A comment on Dutton and Lynn (2015). *Intelligence*, *53*, 81–85. doi:10.1016/j.intell.2015.08.009
- Woods, D. L., Kishiyama, M. M., Yund, E. W., Herron, T. J., Edwards, B., Poliva, O., & ... Reed, B. (2011). Improving digit span assessment of short-term verbal memory. *Journal Of Clinical And Experimental Neuropsychology*, *33*(1), 101 – 111. doi:10.1080/13803395.2010.493149
- Zohar-Shai, B., Tzelgov, J., Karni, A., & Rubinsten, O. (2017). It does exist! A left-to-right spatial–numerical association of response codes (SNARC) effect among native Hebrew

speakers. *Journal of Experimental Psychology: Human Perception and Performance*, 43(4), 719–728. doi:10.1037/xhp0000336