

Université de Montréal

**TRADUCTION FRANÇAISE ET ÉTUDE DE PRÉ-VALIDATION DE LA BATTERIE
D'ÉVALUATION DES FONCTIONS EXÉCUTIVES NIH-EXAMINER CHEZ
L'ENFANT**

Par
Régine Gagnon

Département de psychologie
Faculté des arts et des sciences

Essai doctoral : deuxième dépôt
Présenté en vue de l'obtention du doctorat (Psy.D.)
Option neuropsychologie clinique
Sous la direction de Bruno Gauthier

21 décembre 2018
© Régine Gagnon, 2018

Résumé

Les fonctions exécutives (FE) permettent de réaliser toute action orientée vers un but spécifique requérant un ajustement constant avec les attentes de l'environnement (Luria, 1973). Chez l'enfant, le développement du cortex préfrontal, région associée aux FE, est important. Les FE ont un impact marqué sur plusieurs sphères de son développement. Leur bonne évaluation en neuropsychologie clinique est donc primordiale. Plusieurs outils sont actuellement utilisés sans être validés auprès d'enfants québécois francophones ni offrir de profil complet des FE. Certains ont de faibles qualités psychométriques et sont dispendieux. Cette étude vise à traduire en français la batterie d'évaluation NIH-EXAMINER et à débiter son processus de validation (Kramer *et al.*, 2014). La version originale de cette batterie, disponible en libre accès, démontre d'excellentes propriétés psychométriques et permet d'évaluer la majorité des FE. Les performances à la NIH-EXAMINER traduite, à d'autres outils d'évaluation des FE, du fonctionnement intellectuel et des symptômes d'inattention et d'hyperactivité/impulsivité liés au TDA/H ont été recueillies chez 17 enfants neurotypiques. Les résultats montrent une bonne fidélité en cohérence interne, une bonne validité de construit et de critère, mais demeurent insuffisants pour la validité convergente avec des mesures auto-rapportées de FE dans le quotidien. Des études auprès d'un plus large échantillon et de différents groupes cliniques sont nécessaires pour poursuivre le processus de validation de la version française de la NIH-EXAMINER.

Mots-clés : fonctions exécutives (FE), développement cognitif, validation, psychométrie, évaluation neuropsychologique, enfance.

Abstract

Executive functions (EF) allow any goal-oriented action that requires constant adjustment to environmental expectations (Luria, 1973). In childhood, the development of the prefrontal cortex, a region associated with EF, is particularly important. EF have a strong impact on many areas of development. It is therefore essential to properly evaluate FE in clinical neuropsychology. Several tools are currently used without being validated with Francophone children in Quebec or offering a complete FE profile. Some have poor psychometric qualities and are expensive. This study aims to translate the NIH EXAMINER in French and begin its validation process (Kramer *et al.*, 2014). The original version of this battery, available in open access (free of charge), demonstrates excellent psychometric properties and allows to evaluate the majority of FE. Performance on the translated NIH-EXAMINER, on other assessment tools of FE, as well as on tools assessing intellectual reasoning and symptoms associated with ADHD were collected among 17 neurotypical children. The results show good stability in internal consistency, good criterion and construct validity, but are insufficient in contingent validity for self-reported measures of FE in daily life. Further studies with a larger sample size and different clinical groups are needed to pursue the validation process of the French version of the NIH-EXAMINER.

Keywords: executive functions, cognitive development, validation, psychometry, neuropsychological evaluation, childhood.

Table des matières

Résumé.....	i
Abstract.....	ii
Liste de tableaux.....	v
Liste des abréviations.....	vi
Remerciements.....	viii
Introduction.....	1
Position du problème.....	1
Les fonctions exécutives : le modèle de Miyake <i>et al.</i> (2000).....	2
L’impact des FE sur le développement de l’enfant.....	4
Outils d’évaluation disponibles et limites potentielles.....	6
La batterie d’évaluation des fonctions exécutives NIH-EXAMINER.....	10
Objectifs.....	15
Méthodologie.....	16
Participants.....	16
Procédure.....	16
Mesures.....	18
Analyses statistiques.....	21
Résultats.....	22
Description d’échantillonnage.....	23
Variables contrôles.....	24
Fidélité.....	26
Validité.....	27
Discussion.....	31
Synthèse des résultats.....	31
Critique du cadre théorique et de la conceptualisation de la NIH-EXAMINER.....	32
Critique de l’étude de validation originale de Kramer <i>et al.</i> (2014).....	37
Limites de la présente étude.....	39
Études futures.....	44
Conclusion.....	44

Annexe 1 : Description de la NIH-EXAMINER et données psychométriques de la version originale	i
Annexe 2 : Exemples de stimuli de la NIH-EXAMINER	x
Annexe 3 : Livret de cotation de la NIH-EXAMINER	xiii
Références.....	xlii

Liste de tableaux

Tableau I. Caractéristiques sociodémographiques de l'échantillon.....	23
Tableau II. Scores obtenus à Vocabulaire, Matrices et ARS-IV	24
Tableau III. Scores obtenus aux différentes mesures de FE	25
Tableau IV. Scores obtenus à la NIH-EXAMINER et coefficients de fidélité (cohérence interne).....	26
Tableau V. Intercorrélations des facteurs principaux de FE, de la TNS et du Questionnaire de normes sociales de la NIH-EXAMINER (validité de construit)	28
Tableau VI. Corrélations entre les scores à la NIH-EXAMINER et le BRIEF (validité convergente).....	29
Tableau VII. Corrélations entre la NIH-EXAMINER, les tests cognitifs et l'âge (validité convergente et de critère).....	30

Liste des abréviations

ARS-IV : ADHD Rating Scale – 4th edition

BADS-C : Behavioural Assessment of Dysexecutive Syndrom in Children

BRIEF : Inventaire d'Évaluation Comportementale des Fonctions Exécutives

CVLT-II : California Verbal Learning Test – second edition

DKEFS : Delis–Kaplan Executive Function System

DSM-5 : Diagnostic and statistical manual of mental disorders – 5th edition

FE : fonctions exécutives

RC : rang centile

TDA/H : trouble déficitaire de l'attention avec ou sans hyperactivité

WCST : Wisconsin Card Sorting Test

WISC-IV : Échelle d'intelligence de Wechsler pour enfants et adolescents – Quatrième Édition

WNV: Échelle non verbale d'aptitudes de Wechsler Canadienne

“Try not. Do. Or do not. There is no try.”

- Yoda

Star Wars Episode V – The Empire Strikes Back.

Remerciements

Tout d'abord, j'aimerais remercier Bruno pour ses précieux conseils, son écoute, son enthousiasme devant les projets de recherche et la stimulation des discussions fort intéressantes lors des réunions de laboratoire. Il est primordial pour moi de remercier Laurence, pour l'ensemble de son œuvre ! Depuis mes premières années au Bac, elle a su m'offrir une aide et une amitié inestimables. Je ne pourrai jamais la remercier suffisamment pour tout ce qu'elle a fait. Je remercie également Sofia pour sa rigueur et sa capacité d'organisation hors pairs, Marc-André, pour son efficacité légendaire, Caroline, pour sa disponibilité et sa gentillesse, ainsi que Gabrielle pour ses nombreux conseils et Maude pour son grand cœur. Je remercie également tous les gens du LENÉA pour leurs projets stimulants et leur créativité inspirante.

Je tiens à remercier tous les gens de ma cohorte qui ont traversé le périple du doctorat avec moi, plus particulièrement Édith, Claudia et bien évidemment, Madame Sandra. J'ai eu tellement de chance de vous rencontrer. Je n'y serais pas arrivée sans vous. Je veux aussi remercier certains professeurs qui m'ont profondément marqués et inspirés durant mon parcours : Nadia Gosselin, Nicholas Pesant et Stephan Kennepohl. Vous avez su nourrir encore plus ma passion pour la neuropsychologie, surtout pour l'intervention. Je ne pourrais pas passer sous silence Elaine De Guise qui m'a donné la chance de réaliser ce doctorat, et qui a su accueillir mes inquiétudes et mes préoccupations à plusieurs reprises. Je remercie également Marika Audet-Lapointe, une mentore plus qu'extraordinaire au cœur en or.

Sur une note plus personnelle, j'aimerais remercier Fred de croire que je serai la prochaine Oliver Sacks, de me faire rire quand je doute de moi-même et que je suis découragée et de m'avoir endurée pendant ces derniers temps... Je ne remercierai jamais assez ma mère de

m'avoir toujours encouragée dans mes études, peu importe les choix que j'ai faits et tous les obstacles que j'ai rencontrés. Quand je cessais d'y croire, elle le faisait pour deux ! Je veux aussi remercier mon père qui m'offre son soutien au meilleur de lui-même et qui, je sais, est très fier de moi. Je veux remercier chaleureusement mes « deuxièmes parents », Sylvie et Rosaire d'avoir été aussi présents pour moi. Vous êtes beaucoup plus importants que vous le croyez. Je voudrais remercier Padmé et Mme Chess, parce qu'elles sont des membres de la famille elles aussi et qu'elles ont fait toute la différence dans ma vie à plusieurs occasions. Je souhaite remercier tout particulièrement Gab et Val, des amies comme il ne s'en invente pas et avec qui j'ai l'amitié la plus vraie et la plus solide qui soit. Je remercie également Roxane pour les séances de rédaction interminables et pour le partage mutuel de nos inquiétudes. Je tiens aussi à remercier Emma pour son grand cœur (trop grand parfois) et pour son entrain à toujours réaliser de nouveaux projets avec moi. Finalement, j'aimerais remercier une personne unique qui nous a quittés récemment, mamie Fernande. Cette femme a été une image de courage, de volonté et de persévérance pour plusieurs, particulièrement pour ma mère et moi.

Introduction

Position du problème

Les fonctions exécutives (FE) sont impliquées dans toute action orientée vers un but spécifique et qui requiert un ajustement constant du comportement avec les attentes de l'environnement (Lezak, Howieson, Loring et Fischer, 2004; Luria, 1973; Norman et Shallice, 1986; Pennington et Ozonoff, 1996). Elles permettent de faire des tâches complexes en coordonnant différentes habiletés cognitives (Tranel, 1994). De par leur rôle intégrateur, elles sont impliquées dans toute situation d'apprentissage (Mazeau et Pouhet, 2014).

Anatomiquement, elles sont associées au cortex préfrontal (Luria, 1973), région qui se développe progressivement de l'enfance jusqu'à l'âge adulte (Lussier et Flessas, 2009). Ceci étant dit, à l'enfance, le développement de cette région cérébrale est particulièrement important, et les FE ont un impact majeur sur le développement social, scolaire et affectif (Barkley, 1997). Plusieurs troubles typiquement diagnostiqués à l'enfance sont associés à des atteintes des FE, tels que le syndrome de Gilles de la Tourette, les troubles du spectre de l'autisme et plus particulièrement le trouble du déficit de l'attention / hyperactivité (TDA/H) (Morand-Beaulieu *et al.*, 2017; Nigg, Willcutt, Doyle et Sonuga-Barke, 2005; Pennington et Ozonoff, 1996; Peter, 2010; Sergeant, Geurts et Oosterlaan, 2002; Shallice *et al.*, 2002; Sonuga-Barke, 2003, 2005). La diversité des troubles liés à des déficits de FE et leur importance sur le fonctionnement adaptatif et le développement de l'enfant nécessitent une évaluation neuropsychologique rigoureuse pour fournir des recommandations personnalisées.

Or aucun outil d'évaluation disponibles actuellement, n'a été validé chez une population d'enfants francophones québécois. En plus, peu de ces outils offrent un profil complet des FE, contraignant le clinicien à utiliser plusieurs outils d'évaluation dont les échantillons normatifs

ne sont pas toujours comparables, ont de faibles qualités psychométriques, en plus d'être peu spécifiques ou sensibles aux différentes étiologies possibles. Ensuite, ces outils d'évaluation sont généralement coûteux, dans un contexte actuel de contraintes budgétaires significatives. Ces lacunes peuvent avoir un effet délétère sur la qualité et l'accès aux services offerts. La version originale (en anglais) de la batterie d'évaluation NIH-EXAMINER (Kramer *et al.*, 2014) possède quant à elle d'excellentes propriétés psychométriques, permet d'évaluer l'ensemble des FE et est offerte gratuitement. Ce projet vise ainsi à traduire en français et à décrire les données psychométriques de la batterie d'évaluation des FE NIH-EXAMINER.

Les fonctions exécutives : le modèle de Miyake *et al.* (2000)

Bon nombre d'auteurs ont tenté de définir le construit complexe des FE (Anderson, 2002; Baddeley, 1986, 2000; Barkley, 1997; Lezak, 1995; Miyake *et al.*, 2000; Norman et Shallice, 1986; Shallice et Burgess, 1996; Stuss, Shallice, Alexander et Picton, 1995; Zelazo, Carter, Reznick et Frye, 1997). Or, encore à ce jour, aucun modèle ne fait l'unanimité, tant au niveau des composantes que de leur rôle exact ou même de leur définition (Anderson, Jacobs et Anderson, 2010; Jurado et Rosselli, 2007; Packwood, Hodgetts et Tremblay, 2011). Le modèle de Miyake *et al.* (2000) à trois composantes étant le cadre conceptuel de la batterie d'évaluation des FE NIH-EXAMINER, il sera le modèle de référence utilisé dans ce présent projet de validation en langue française.¹

Afin de pallier le manque de consensus dans la littérature concernant l'organisation des

¹ Pour une revue des FE plus exhaustive, critique et pertinente au point de vue clinique, il est suggéré de se référer aux publications et ouvrages suivants : (Anderson *et al.*, 2010; Barkley, 2012; Luria, 1973; Suchy, 2009, 2015).

FE, l'étude de Miyake *et al.* (2000) avait pour premier but de déterminer si les FE représentaient un concept unitaire ou fragmenté en différentes FE. Puis, un second but était de déterminer les liens unissant les différentes FE.

Une des FE documentée est la flexibilité cognitive « *shifting of mental set* », définie comme la capacité à s'adapter aux changements induits par l'environnement pour atteindre un but spécifique (Lewis et Carpendale, 2009; Zelazo *et al.*, 1997). Une seconde composante est le monitoring, nommé aussi mise à jour ou « *monitoring and updating information in working memory* », est quant à lui défini comme l'habileté à maintenir, à manipuler et à rejeter l'information non-pertinente des représentations mentales en mémoire de travail pour atteindre le but souhaité (Miyake *et al.*, 2000; Morris et Jones, 1990). Une troisième composante est l'inhibition, décrite comme la capacité à choisir de maintenir une réponse non-automatique et utile à l'atteinte d'un but (Miyake *et al.*, 2000). Selon ce modèle, lors de tâches plus complexes, les trois sous-composantes seraient sollicitées (Miyake *et al.*, 2000). Elles interviendraient dans la coordination de plusieurs tâches en même temps (double tâche) et en planification, pour établir et mettre en œuvre une séquence d'événements organisés dans le temps selon un objectif (Anderson *et al.*, 2010).

Miyake *et al.* (2000) ont choisi ces trois composantes de FE parce qu'elles leur semblaient plus élémentaires que des FE plus complexes comme la planification, que leur définition est facile à opérationnaliser, qu'elles sont associées à des outils d'évaluation simples et bien étudiés et qu'elles sont probablement impliquées dans la réalisation de tâches plus complexes. Ils ont administré à 137 adultes neurotypiques divers tests d'évaluation des composantes de FE et présentant une variabilité dans la nature des tâches (par ex. bimodale visuelle / auditive) ainsi que des tests plus complexes. Les analyses factorielles exploratoires

conduites a posteriori ont permis de confirmer l'hypothèse d'un modèle à trois sous-composantes de FE et d'en démontrer les interrelations (coefficients de corrélation entre 0,42 et 0,63). Ils ont donc statué que les FE ne représentaient pas un construit strictement fragmenté en trois sous-composantes et qu'il y aurait une composante de FE globale les reliant.

En résumé, le modèle de Miyake *et al.* (2000) offre une conceptualisation et une organisation des FE particulièrement parcimonieuses en identifiant seulement trois sous-composantes principales de FE, tout en prenant en compte le rôle global de régulation cognitive des FE. De plus, la méthode d'analyses confirmatoires évite d'obtenir un nombre trop élevé de composantes et de multiplier les termes associés. Or, procéder ainsi ne permet pas de prendre en compte l'ensemble des facteurs impliqués et pourrait mener à une trop grande simplification du construit des FE. Ce modèle demeure cependant un bon point de départ dans l'élaboration d'un outil d'évaluation des FE comme la NIH-EXAMINER.

Une étude réalisée auprès de 108 enfants âgés de 8 à 13 ans a permis de confirmer ce modèle en montrant des résultats similaires, qui appuient un modèle à trois facteurs inter-reliés : la flexibilité, la mémoire de travail (fonction similaire à la mise à jour) et l'inhibition (Lehto, Juujärvi, Kooistra et Pulkkinen, 2003). Les auteurs ont également procédé par analyses factorielles confirmatoires et exploratoires a posteriori pour contre-valider leurs résultats. L'étude de Lehto *et al.* (2003) permet donc de soutenir la pertinence du modèle de Miyake *et al.* (2000) chez les enfants.

L'impact des FE sur le développement de l'enfant

Chez l'enfant, les FE se développent progressivement (Lussier et Flessas, 2009) et influencent grandement ses apprentissages et comportements (Barkley, 1997; Mazeau et Pouhet, 2014), notamment sur les plans social, scolaire et émotionnel (Barkley, 1997).

Les FE influenceraient certaines habiletés sociales telles que la théorie de l'esprit, la communication sociale et le jugement moral. La théorie de l'esprit réfère à la capacité d'être conscient de son état mental ou de celui des autres (Moriguchi, 2014). Elle permet à l'enfant de concevoir que ses croyances peuvent être biaisées et fausses. Lorsque l'enfant a un fonctionnement exécutif adéquat, il serait plus en mesure d'observer et de porter un jugement sur son propre état mental (monitorage) et celui des autres. Les FE permettraient également une meilleure compréhension et respect des normes sociales, et faciliteraient le jugement moral, c.-à-d. la capacité à appliquer des règles multidimensionnelles et des standards pour évaluer les intentions et la probité de ses conduites et de celles d'autrui (Bandura, 2014; Moriguchi, 2014). L'ajustement comportemental de l'enfant aux normes sociales véhiculées par son environnement serait meilleur lorsque les FE sont intactes (Rueda et Paz-Alonzo, 2013).

Les FE ont également un impact considérable sur le développement scolaire de l'enfant (Barkley, 1997). L'acquisition de FE plus complexes pendant l'enfance, telles que la sélection et la coordination de fonctions cognitives appropriées pour atteindre un but, permet un meilleur rendement scolaire, et la mémoire de travail et l'inhibition influenceraient positivement ce rendement, indépendamment de l'âge (Best, Miller et Naglieri, 2011). De plus, l'intégrité de FE telles que la planification, le monitorage et la flexibilité cognitive auraient un impact positif sur lui (Blair et Diamond, 2008). Au niveau du développement affectif, l'inhibition et le contrôle attentionnel faciliteraient la régulation émotionnelle (Carlson et Wang, 2007; Rueda et Paz-Alonzo, 2013; Simonds, Kieras, Rueda et Rothbart, 2007). Par exemple, l'enfant apprendrait à exprimer ses émotions aux autres en inhibant une manière inappropriée de le faire (par ex. crier contre les autres lorsqu'il ressent de la colère).

En résumé, l'intégrité des FE chez l'enfant interagirait étroitement avec la qualité de son développement social. Le rendement scolaire et la gestion émotionnelle chez les enfants seraient également facilités par un développement sain des FE.

Outils d'évaluation disponibles et limites potentielles

Plusieurs outils standardisés existent déjà et sont fréquemment utilisés en neuropsychologie clinique pour évaluer différentes FE telles que l'inhibition, la flexibilité, la fluidité et la planification. La présente section présentera les tests et batteries d'évaluation des FE les plus communément utilisés actuellement en neuropsychologie clinique chez les enfants. Ces outils demeurent pertinents pour évaluer l'évaluation des FE, mais présentent certaines lacunes.

La *Delis-Kaplan Executive Function System (DKEFS)* (Delis, Kaplan et Kramer, 2001) a été conceptualisée aux États-Unis pour évaluer la population américaine âgée de 8 à 89 ans. Cette batterie évalue plusieurs FE (c.-à-d. la fluidité, l'inhibition, la flexibilité et la planification), fournit des scores pour plusieurs tranches d'âges. Les sous-tests peuvent être réalisés individuellement dans n'importe quel ordre, et la durée totale de passation est d'environ 90 minutes (Strauss, Sherman et Spreen, 2006). L'élaboration de la *DKEFS* a été conçue en s'inspirant de modèles de FE qui permettent de mieux comprendre les dysfonctions associées aux lobes frontaux (Baddeley, 1986; Luria, 1980; Shallice et Burgess, 1996) et selon une approche basée sur les processus. Ainsi, les auteurs (Delis *et al.*, 2001) ont inclus des scores complémentaires (p.ex. erreurs commises, bris de règles, mesures par intervalles) dans le but d'évaluer l'aspect multifactoriel des FE. Malgré l'apport de cette approche, la *DKEFS* présente des lacunes pour l'évaluation des FE. Par exemple, plusieurs tests ne sont pas disponibles en français, ni aucune norme québécoise francophone, et la traduction des items

des tests n'a pas été validée empiriquement. La *DKEFS* ne peut donc être utilisée que partiellement au Québec et n'est pas en mesure d'établir un profil valide des FE auprès d'une population francophone. De plus, les tests Fluidité verbale, Tour et Traçage de séquences possèdent une faible fidélité en cohérence interne et en équivalence par test-retest (entre 0,59 et 0,69 respectivement). Pour les tests dont une version alternative existe, c.-à-d. une version modifiée qui mesure le même construit (Test du tri, les 20 questions et Fluidité verbale), les corrélations sont faibles (0,44 pour la condition 4 de Fluidité verbale). Ceci suggère une faible fidélité en équivalence par méthode des versions alternatives.

La *Behavioural Assessment of Dysexecutive Syndrom in Children (BADS-C)* (Emslie, Burden, Nimmo-Smith, Wilson et Wilson, 2003) est une batterie d'Angleterre visant à évaluer de manière écologique différentes FE et comportements associés tels que la planification, la flexibilité et les persévérations, la résolution de problèmes, l'utilisation de la rétroaction, l'impulsivité, le respect des consignes et le séquençage. La *BADS-C* s'appuie sur deux modèles de FE largement cités dans la littérature, soit le modèle de mémoire de travail de Baddeley et Hitch (1974) et le modèle du système de contrôle attentionnel (*SAS : Supervisory Attentional System*) de Norman et Shallice (1986); (Shallice, 1982). La durée de passation est d'environ 35 à 45 minutes. Bien que la batterie fournisse un profil assez complet des FE, les indices de fidélité test-retest, de validité convergente et de validité de critère se sont avérés plutôt faibles (Baron, 2007). De plus, aucune adaptation francophone validée ou normée de la *BADS-C* n'est disponible.

Le *Wisconsin Card Sorting Test (WCST)* (Heaton, Chelune, Talley, Kay et Curtiss, 1993) est utilisé pour évaluer principalement la flexibilité cognitive, et sa durée est de 15 à 30 minutes (Strauss *et al.*, 2006). Ce test, conçu aux États-Unis, avait pour but de mesurer le

raisonnement abstrait et les stratégies de résolution de problèmes. Le *WCST* est une mesure des FE selon la conceptualisation de Luria (1973) et de Shallice (1982) puisqu'il vise à évaluer la capacité à développer une stratégie appropriée pour atteindre un but lorsque les conditions changent. Ce test ne fournit pas un profil de différentes composantes de FE, mais a gagné en popularité depuis l'élaboration d'une passation standardisée (Heaton, 1981). En raison du fait qu'il n'est plus possible de réaliser le test à nouveau si la personne parvient à trouver la stratégie à employer, la fidélité test-retest du *WCST* est faible (jusqu'à 0,12) (Strauss *et al.*, 2006). En plus d'évaluer l'abstraction et la mise en place de stratégies, le *WCST* évalue également le maintien de l'engagement et de la vigilance dans une tâche, la capacité à prendre en compte la rétroaction, le traitement visuel de base et la capacité à moduler son état mental et motivationnel. Un faible score au *WCST* peut être lié à plusieurs facteurs impliqués dans les FE, autre que la présence d'un déficit spécifique de flexibilité cognitive. Bien que le *WCST* demeure pertinent pour l'évaluation des FE, il est insuffisant pour dresser un profil exécutif complet, et les scores obtenus doivent être interprétés avec précaution.

Le Test d'Évaluation de l'Attention Chez l'Enfant (TEA-Ch) (Manly *et al.*, 2006) a été développé en Angleterre, et une adaptation francophone a été réalisée en France. La structure de la TEA-Ch est basée sur une conceptualisation fractionnée de l'attention selon le modèle de (Posner et Petersen, 1990) et selon les données recueillies chez des adultes à l'aide du Test d'Évaluation de l'Attention (TEA) (Robertson, Ward, Ridgeway et Nimmo-Smith, 1994). Cette batterie permet d'évaluer des systèmes attentionnels divisés en trois facteurs : l'attention sélective, l'attention soutenue et le contrôle attentionnel/la flexibilité en plus d'évaluer d'autres FE telles que l'inhibition et la mémoire de travail. La version originale de la TEA-Ch a été validée auprès de 293 enfants australiens. L'adaptation française a été validée auprès de

379 enfants de France. Les données d'étalonnage recueillies lors de l'adaptation française ne suivent pas une courbe normale et les scores obtenus sont donc des pourcentages cumulés plutôt que des percentiles. Ceci rend plus difficile l'interprétation des scores, car ils ne renseignent pas sur le rang réel du score de l'enfant par rapport à la norme. De plus, peu de données concernant la validité sont rapportées dans l'adaptation française, et la fidélité test-retest est insatisfaisante pour le sous-test des Petits hommes verts entre les deux versions parallèles, c.-à-d. mesurant exactement le même construit.

Le Bilan neuropsychologique de l'enfant – Seconde édition (NEPSY-II) (Korkman, Kirk et Kemp, 2014) a été élaboré aux États-Unis et adapté en France. Cette batterie est basée sur une conceptualisation dynamique et développementale des FE inspirée de l'approche qualitative de Luria (1973, 1980). La NEPSY-II vise à évaluer les sous-composantes fondamentales ainsi que les sous-composantes plus complexes de chaque domaine de FE. Les sous-tests Attention auditive et Réponses associées, Catégorisation, Fluidité de dessins, Horloges, Inhibition et Statue permettent d'évaluer différentes FE telles que l'inhibition, la flexibilité, la fluidité et la planification ainsi que le contrôle attentionnel. Les qualités psychométriques de la NEPSY-II en termes de fidélité et de validité sont satisfaisantes. Toutefois, aucune norme française n'est disponible. Ce sont donc les données de l'échantillon américain qui servent de point de repère pour l'interprétation des scores. La NEPSY-II demeure une batterie pertinente pour l'évaluation des FE et du contrôle attentionnel, mais limitée car non adaptée à la population québécoise francophone.

En résumé, plusieurs outils standardisés sont fréquemment utilisés en neuropsychologie clinique et demeurent pertinents pour évaluer les FE. Cependant, ces outils peuvent poser différents problèmes comme l'absence de normes et de validation chez une

population d'enfants québécois francophones. Les normes disponibles, provenant souvent des États-Unis ou de l'Europe, ne sont pas représentatives de la population québécoise, ce qui rend difficile l'interprétation des scores obtenus. De plus, plusieurs outils n'offrent pas un profil complet des FE ou ont de faibles qualités psychométriques. Enfin, le coût élevé de ces outils n'est pas à négliger, ce qui contribue au prix élevé des évaluations et affecte l'accès aux services. Un outil d'évaluation des FE francophone moins long, plus complet, moins dispendieux et adéquat au niveau psychométrique apparaît donc nécessaire.

La batterie d'évaluation des fonctions exécutives NIH-EXAMINER

En 2014, des chercheurs du *National Institute of Health (NIH)* ont élaboré la NIH-EXAMINER (Kramer *et al.*, 2014), une batterie d'évaluation des FE libre d'accès qui permet de dresser un profil complet des FE et dont le temps d'administration est court (45 minutes). Comme plusieurs sous-tests de la batterie sont informatisés, la NIH-EXAMINER permet de générer une grande variété de scores à chacun des sous-tests tout en diminuant les erreurs de passation et de cotation des expérimentateurs. La NIH-EXAMINER est disponible en trois versions (A, B et C), offerte en trois langues (anglais, espagnol et hébreu) et inclut une version adaptée à l'évaluation des enfants.

La vocation primaire de la NIH-EXAMINER est d'enrichir et d'uniformiser les pratiques d'évaluation des FE en recherche. Outil d'évaluation des FE basé sur le modèle théorique bien établi de Miyake *et al.* (2000), elle vise à offrir rapidement un profil général de plusieurs FE différentes. En recherche, une méthode d'évaluation uniformisée des FE favorise une meilleure opérationnalisation du construit des FE et une meilleure réplique des études scientifiques. Par ailleurs, en clinique, la NIH-EXAMINER pourrait s'avérer pertinente. L'uniformisation des méthodes d'évaluation en clinique est utile, puisqu'elle permet un

consensus entre cliniciens sur les méthodes d'évaluation et d'interprétation du profil exécutif, sur les diagnostics associés à ces profils et sur les recommandations fournies par la suite.

Brève description de la batterie NIH-EXAMINER. La version pour enfants de la NIH-EXAMINER comprend différents sous-tests permettant d'évaluer les trois composantes du modèle de Miyake *et al.* (2000), soit la flexibilité, l'inhibition et la mémoire de travail, équivalente à la mise à jour selon Kramer *et al.* (2014). D'abord, la NIH-EXAMINER comprend un sous-test évaluant la flexibilité cognitive et deux sous-tests évaluant l'inhibition. La flexibilité est évaluée par le Test d'appariement où l'enfant doit appairer selon la forme ou la couleur des stimuli qui lui sont présentés à l'écran. La consigne change en cours de tâche, et l'enfant doit adapter ses réponses en conséquence. L'inhibition est évaluée via le *Flanker* et le *CPT*. Le *Flanker* consiste à présenter cinq poissons sur un écran d'ordinateur à l'enfant qui doit identifier dans quelle direction nage le poisson du centre sans se laisser influencer par la direction de nage des quatre poissons environnants.

Quant à la mise à jour, Kramer *et al.* (2014) ont décidé de remplacer cette fonction par la mémoire de travail qu'ils estiment un concept permettant d'englober le traitement en mémoire à court terme de l'information et son monitoring. La version pour enfants de la NIH-EXAMINER inclut donc deux sous-tests évaluant la mémoire de travail : Comptage de cercles et *N-Back*. Le Comptage de cercles consiste à compter les cercles bleus qui apparaissent à l'écran parmi des éléments de distraction. L'enfant doit maintenir en mémoire le nombre de cercles bleus comptés à chaque item et les rappeler dans l'ordre à la fin de l'essai. Chaque essai comprend de plus en plus d'items, et donc de plus en plus de nombres à retenir. Le *N-Back* consiste à comparer la localisation d'une forme présentée à celle d'une autre forme présentée juste avant sur l'écran d'ordinateur.

Afin que leur batterie fournisse un profil exécutif plus exhaustif et qu'elle s'appuie sur les pratiques courantes d'évaluation, les auteurs ont inclus des sous-tests visant à évaluer la fluidité, la planification et l'autocritique (*insight*). La fluidité est la capacité à produire un maximum de réponses, verbales ou graphiques, en un temps limité (Delis *et al.*, 2001). Les concepteurs de la NIH-EXAMINER (Kramer *et al.*, 2014) ont inclus un sous-test de fluidité verbale catégorielle, car il s'agit d'une fonction fréquemment associée aux FE dans la littérature.

La planification consiste à concevoir et à mettre en œuvre une séquence d'évènements au niveau temporel, dans le but d'atteindre un objectif précis (Anderson *et al.*, 2010). La planification est souvent décrite comme une sous-composante importante des FE dans la littérature (Lezak, 1995; Norman et Shallice, 1986; Zelazo *et al.*, 1997). Miyake *et al.* (2000) mentionnent d'ailleurs qu'il s'agit d'une fonction plus complexe issue des trois sous-composantes principales (c.-à-d. flexibilité, monitoring et inhibition). Ainsi, les concepteurs de la NIH-EXAMINER (Kramer *et al.*, 2014) ont inclus le Test non-structuré pour évaluer la planification. L'enfant doit choisir, parmi trois cahiers, différentes tâches à réaliser, telles que relier des points, biffer des cibles, compléter des labyrinthes ou reproduire des symboles en priorisant celles qui lui permettent de remporter le plus de points le plus rapidement possible.

L'autocritique est la capacité de s'auto-observer et de se connaître soi-même (Markova et Berrios, 1992). La NIH-EXAMINER (Kramer *et al.*, 2014) contient une mesure qualitative visant à l'évaluer. Après la tâche de Fluidité, l'enfant doit évaluer et comparer sa performance à l'aide d'une courbe normale qui lui est expliquée au préalable. La manière dont l'enfant interagit avec les autres et se comporte au quotidien reflète sa capacité de régulation comportementale. Selon Kramer *et al.* (2014), il s'agit d'une mesure non cognitive des FE et

c'est pourquoi ils ont tenu à inclure un questionnaire de normes sociales rempli par l'enfant et une échelle comportementale remplie par l'évaluateur. De plus amples détails sur les sous-tests et des exemples de stimuli sont présentés en Annexes.

Étude de validation de la version originale de la NIH-EXAMINER. La NIH-EXAMINER a été administrée à des enfants par les membres de laboratoires de recherche de la University of California, San Francisco (UCSF), en collaboration avec neuf milieux de recherche incluant des universités et des centres hospitaliers : Boston Children's Hospital, Case Western Reserve, University of Nebraska-Lincoln et University of South Carolina. Cette procédure visait à recueillir les données d'enfants d'âges différents (3 à 17 ans), provenant de régions géographiques différentes, appartenant à différents groupes ethniques et cliniques. Les groupes cliniques étaient constitués d'enfants présentant un TDA/H ($n = 34$), un très faible poids à la naissance ($n = 72$) et une anémie falciforme ($n = 34$). Les données des enfants neurotypiques ont été recueillies suivant la même procédure que celle décrite ci-dessus pour assurer la diversité de l'échantillonnage ($n = 337$)².

Le détail des indices psychométriques disponibles tirés de la validation de la version originale est présenté à l'Annexe 1 et tiré du manuel de l'examineur produit par Kramer *et al.* (2014) fourni avec le reste du matériel de la batterie (logiciel d'installation PsychoPy, protocoles, vidéos explicatifs, etc.). Certains indices ne sont disponibles que pour la version adulte dans le manuel de l'examineur ; ils seront rapportés le cas échéant.

Les qualités psychométriques font référence à la fidélité du test, c.-à-d. la constance d'un test, et à la validité, qui détermine si le test mesure bien ce qu'il est censé mesurer

² Manuel de l'examineur fourni par Kramer et al. (2014).

(Hogan, 2017). Les qualités psychométriques de la version originale de la NIH-EXAMINER s'avèrent satisfaisantes autant au niveau du score global de FE que des facteurs et des tests individuels et proviennent de larges échantillons d'enfants et d'adultes. Effectivement, en ce qui a trait à la fidélité, la NIH-EXAMINER possède une bonne cohérence interne par méthode de covariances (par ex. alpha de Cronbach), une bonne fidélité test-retest et un accord inter-juges adéquat³. Un coefficient de fidélité de 0,90 est jugé élevé, de 0,80 est jugé modérément élevé, de 0,70 est jugé faible et en-deçà de 0,60, il est jugé inacceptable (Hogan, 2017).

La NIH-EXAMINER possède une validité de construit, une validité convergente et une validité divergente satisfaisantes, ainsi qu'une validité de critère adéquate permettant de distinguer les participants contrôles des participants cliniques ainsi qu'un effet de l'âge sur certains scores obtenus⁴. Tel que rapporté dans Hogan (2017), un coefficient de validité de 0,20 est jugé faible, de 0,50 est jugé moyen et de 0,80 est jugé élevé (Cohen, 1988).

³ La cohérence interne fait référence aux relations qui existent entre les items ou ensembles d'items à l'intérieur d'un même test. La fidélité test-retest consiste à administrer le même test aux mêmes personnes à différents temps de mesure et la fidélité inter-juges permet d'évaluer les variations attribuables aux personnes qui corrigent le test.

⁴ La validité de construit permet d'évaluer si le test mesure bien le construit théorique ou phénomène qu'il est censé mesurer, alors que la validité de critère évalue la relation entre le test et un critère externe (Hogan, 2017). Un type de validité de critère est la validité concomitante, qui établit la relation entre le test et un critère externe mesuré au même moment. La validité de critère concomitante peut être réalisée par méthode de groupes contrastes (par ex. groupes cliniques et groupes contrôles), ou avec un critère externe tel que l'âge (Hogan, 2017). Certains auteurs définissent la validité convergente et divergente comme un type de validité de construit (Campbell et Fiske, 1959; Cronbach et Meehl, 1955) alors que d'autres (Hogan, 2017) la définissent comme un type de validité de critère. Peu importe la classification, la validité convergente permet d'évaluer la relation entre deux tests présumés évaluer le même construit, et la validité divergente, la relation entre deux tests présumés évaluer des construits opposés (Hogan, 2017).

En résumé, la version originale de la NIH-EXAMINER montre des indices de fidélité et de validité satisfaisants. Cette batterie constitue un outil pertinent pour l'évaluation des FE chez les enfants, mais aucune version en français n'est disponible à ce jour.

Objectifs

Cet essai doctoral vise à présenter les résultats préliminaires du processus de traduction et de validation de la version en français de la batterie d'évaluation des FE NIH-EXAMINER. Il décrit la méthode de traduction employée et les performances obtenues par des enfants québécois francophones aux différentes mesures de la NIH-EXAMINER. Il vise à fournir des données préliminaires sur les propriétés psychométriques de la version traduite quant à sa validité et à sa fidélité. Ces données psychométriques pourront être comparées à celles de l'étude de validation originale (Kramer *et al.*, 2014) afin d'explorer la pertinence d'un tel outil auprès des enfants québécois francophones.

La traduction et la validation de cette batterie s'avèrent hautement pertinentes. Détenir un outil adapté à la population d'appartenance de l'enfant évalué permet une meilleure comparabilité des scores. Par exemple, lorsque les scores obtenus par un enfant québécois francophone sont comparés à ceux obtenus par une population d'enfants provenant des États-Unis ou d'Europe, la marge d'erreur est plus grande. Les différences démographiques, socio-économiques, scolaires, langagières et culturelles entre l'échantillon normatif et l'enfant évalué peuvent contaminer la précision du score obtenu. Il a été montré que les environnements génétique, périnatal et culturel affectent l'organisation du cerveau et la façon dont les nouveaux apprentissages sont traités (Kennepohl, 1999). Il a aussi été démontré que les valeurs culturelles influencent le rendement aux tests psychométriques (Ardila, 2005). Ces différences entre les populations créent de l'interférence et peuvent nuire à la validité des

scores et, par conséquent, du profil cognitif. Ces répercussions de l'absence d'outil adapté aux enfants québécois francophones sont non négligeables.

Méthodologie

Participants

L'échantillon final était constitué de 17 enfants âgés de 8 à 12 ans (M : 10,16 ; ÉT : 1,11), un participant ayant eu un résultat déficitaire (-2 ÉT) à une mesure contrôle évaluant le niveau intellectuel. Les critères d'exclusion comprenaient la présence d'un trouble neurologique ou neurodéveloppemental ou d'une déficience intellectuelle. Ils étaient vérifiés lors de l'explication des conditions de participation à l'étude par des tests cognitifs (voir ci-dessous la section Outils d'évaluation; *Vocabulaire et Matrices*). La présence de symptômes d'inattention et d'hyperactivité/impulsivité était vérifiée par un questionnaire de dépistage du TDA/H (voir Questionnaires aux parents; *ARS-IV*).

Procédure

Les enfants ont été recrutés dans une école primaire, un camp de jour et via internet (Kijiji, Facebook). Après remise de leurs coordonnées aux chercheurs par le biais du formulaire de sollicitation, leurs parents ont été contactés par courriel ou par téléphone pour connaître leur intérêt à participer à l'étude et expliquer les modalités de l'étude (durée, compensation, nature des tâches effectuées, etc.). Le formulaire de consentement leur a été transmis par courriel. Les parents étaient invités à communiquer par courriel avec l'équipe pour établir un moment de rencontre convenant à leur horaire et pour recevoir de l'information supplémentaire si nécessaire.

Traduction de la NIH-EXAMINER. La traduction en français de la NIH-EXAMINER a suivi la méthode inversée de Vallerand (1989) pour s'assurer de la comparabilité avec la version originale anglophone. Elle a d'abord été réalisée par un membre bilingue de l'équipe de recherche. Puis, une retraduction du français vers l'anglais a été effectuée par un autre membre bilingue de l'équipe. Les deux membres de l'équipe se sont rencontrés pour comparer les versions anglaises et se sont entendus sur l'ajustement du libellé de certains items. Un troisième membre de l'équipe, aveugle au processus de traduction, a ensuite donné son avis pour un dernier ajustement. Finalement, les deux membres de l'équipe traducteurs se sont rencontrés à nouveau pour s'entendre sur la version française finale administrée aux participants. Bien qu'ils aient rapporté un niveau d'aisance en anglais suffisant pour bien comprendre les items des sous-tests, leur bilinguisme n'a pas été évalué formellement. Le manuel d'instructions, de formation et les manuels à l'usage de l'expérimentateur ont été traduits de l'anglais au français sans suivre la procédure de Vallerand, puisque les participants ne sont pas en contact direct avec ce matériel.

Évaluation des enfants. Chaque enfant a été rencontré environ deux heures pour l'évaluation des FE avec la batterie NIH-EXAMINER et pour d'autres épreuves évaluant la mémoire de travail (Séquences de chiffres et Empan spatial), l'inhibition et la flexibilité cognitive (Interférence couleurs-mots), la fluidité (*5-Point Test*), ainsi que pour une évaluation abrégée du raisonnement perceptif et de la compréhension verbale (Vocabulaire, Matrices ; *WISC-IV*). Deux questionnaires d'évaluation comportementale, le *ADHD Rating Scale – 4th edition (ARS-IV)* (Dupaul, Power, Anastopoulos et Reid, 1998) et le *BRIEF*, ainsi qu'un questionnaire de données sociodémographiques ont été remplis par le parent.

La NIH-EXAMINER a été exécutée sur un PC avec Windows 8 par le logiciel PsychoPy (Version V1.73.02). Des instructions et stimuli sous forme papier-crayon étaient également utilisés pour certains tests (voir Annexe 3 pour le cahier de passation).

Mesures

Séquences de chiffres (WISC-IV) (Wechsler, 2004). L'enfant doit répéter des séquences de chiffres présentés oralement en ordre direct puis inverse, ce qui permet d'évaluer la capacité d'encodage et de manipulation de l'information verbale en mémoire de travail. Les coefficients de fidélité pour les tests Séquences de chiffres sont bons ($r = 0,86$ et $r = 0,71$, respectivement). La validité est également bonne, le test Séquences de chiffres étant corrélé fortement ($r = 0,80$) à l'indice de mémoire de travail (IMT) de la batterie complète. Les mesures retenues pour la présente étude sont le nombre d'items en ordre inverse réussis (scores bruts).

Empan spatial (WNV) (Wechsler et Naglieri, 2006). Le test Empan spatial évalue la capacité de l'enfant à maintenir en mémoire une séquence visuelle, à la manipuler pour la mettre dans un certain ordre, puis à la répéter. Des blocs identiques sont disposés de manière non systématique sur une planche de 8 x 11 pouces. Le test comprend deux conditions, en ordre direct puis indirect. L'évaluateur touche les blocs dans un ordre précis, puis l'enfant doit reproduire la séquence dans l'ordre direct. À la deuxième séquence présentée, l'enfant doit la reproduire dans l'ordre inverse. Les coefficients de fidélité pour la condition en ordre direct ($r = 0,76$) et en ordre inverse ($r = 0,82$) sont bons. La validité est acceptable, Empan spatial étant corrélé moyennement ($r = 0,48$) au score obtenu à la batterie *WNV* complète chez des personnes âgées de 8 à 21 ans. Les mesures retenues sont le nombre d'items ordre inverse

réussis. Pour cette étude, une version informatisée sur ordinateur a été administrée aux participants.

Interférence couleurs-mots (DKEFS) (Delis *et al.*, 2001). Le sous-test Interférence couleurs-mots s'inspire de la procédure de Stroop (1935) et mesure l'inhibition. Le test comprend deux conditions de base (dénomination de couleurs et lecture de mots simples). À la troisième condition, celle d'interférence, l'enfant doit inhiber une réponse automatique de lecture de mots pour nommer la couleur de l'encre utilisée pour écrire les mots. Enfin, à la dernière condition, qui évalue la flexibilité cognitive, l'enfant doit alterner entre l'inhibition d'une réponse automatique et la lecture du mot. La fidélité test-retest pour ce test est excellente ($r = 0,90$). La vitesse d'exécution ainsi que le nombre d'erreurs corrigées et non corrigées sont collectés pour chaque condition. Ce sont ces mesures aux conditions d'inhibition et de flexibilité qui sont retenues pour la présente étude (scores bruts).

5-Point Test (Regard, Strauss et Knapp, 1982). La Fluidité graphique est un test de raisonnement non verbal considéré comme l'analogue visuel de la fluidité verbale. L'enfant doit produire en cinq minutes le plus grand nombre de graphiques non symboliques différents en reliant des points à l'aide de lignes droites dans une matrice de 40 rectangles, dont chacun comprend cinq points disposés comme sur un dé. Les mesures d'intérêt sont le nombre total de graphiques (évaluant l'activation), le nombre d'erreurs de répétitions (évaluant l'autorégulation) et le nombre de stratégies (par ex. inversion, ajout de lignes, évaluant la planification). La version papier-crayon de ce test inclut des normes pour les enfants de 1^{re}, 3^e, 5^e et 7^e années. Lors de la présente étude, une version alternative du *5-point Test*, sur tablette numérique, a été administrée aux participants.

Vocabulaire et Matrices (WISC-IV) (Wechsler, 2004). Les sous-tests Vocabulaire et

Matrices permettent d'obtenir un estimé du fonctionnement intellectuel, plus précisément de l'indice de compréhension verbale (ICV) et de l'indice de raisonnement perceptif (IRP). Ils sont corrélés modérément à l'indice de quotient intellectuel global (QIG), à 0,68 et à 0,55 respectivement. Ils étaient utilisés afin d'exclure les participants ayant un score déficitaire (-2 ÉT) à l'un des deux sous-tests, en supposant que ces participants auraient également obtenu un QIG déficitaire. Le sous-test Vocabulaire évalue l'habileté de l'enfant à définir des mots présentés oralement par l'expérimentateur. Le sous-test Matrices évalue l'habileté à analyser des matrices de graphiques incomplètes et à identifier les patrons qui les complètent logiquement. Le score standardisé de bonnes réponses obtenues à chaque test a été utilisé.

ARS-IV (Dupaul *et al.*, 1998). Dans le but d'exclure les participants présentant un seuil clinique de symptômes associés au TDA/H, le questionnaire *ARS-IV* version parent a été administré pour évaluer la présence d'un TDA/H chez leur enfant. Il comprend la liste des neuf symptômes d'inattention et des neuf symptômes d'hyperactivité/impulsivité du TDA/H proposés par le *Diagnostic and Statistical Manual of Mental Disorders – 5th edition (DSM-5)* ; *American Psychiatric Association (APA)*, 2013. Le parent répondant doit identifier dans quelle mesure le comportement cité correspond au comportement de l'enfant sur une échelle Likert de 0 à 3 (jamais ou rarement ; quelques fois ; souvent ; très souvent). Les coefficients de cohérence interne de cet instrument sont élevés (0,86 à 0,92). De même, les indices de fidélité test-retest sont adéquats (0,78 à 0,90). Les mesures retenues pour cette étude sont les scores en rangs centiles (RC) aux échelles d'inattention, d'hyperactivité/impulsivité et du score total de symptômes.

BRIEF (Gioia, Isquith, Guy et Kenworthy, 2000). Le questionnaire *BRIEF* a pour objectif d'évaluer des comportements quotidiens associés aux FE. Il est constitué de huit sous-

échelles qui font référence à l'inhibition, la flexibilité, le contrôle émotionnel, l'initiation, la mémoire de travail, la planification, l'organisation du matériel et au monitoring. Deux indices sont ensuite obtenus : l'index de régulation comportementale, comprenant les sous-échelles d'inhibition, de flexibilité et de contrôle émotionnel, et l'index de métacognition, aussi appelé index de régulation cognitive, comprenant les sous-échelles d'initiation, de mémoire de travail, de planification, d'organisation du matériel et de monitoring. Un indice global du fonctionnement exécutif est obtenu comprenant les indices de régulation comportementale et de métacognition. Ce questionnaire de 86 items est rempli par le parent selon une échelle en trois points (jamais ; parfois ; souvent). L'alpha de Cronbach pour la version parent est élevé ($r = 0,80$). Les mesures utilisées sont le score global, les scores des deux indices et les scores de plusieurs sous-échelles (inhibition, flexibilité, contrôle émotionnel, mémoire de travail et planification).

Analyses statistiques

En plus des scores de performance obtenus à chacune des tâches, des scores factoriels sur les trois domaines exécutifs (Fluidité, Contrôle cognitif et Mémoire de travail) peuvent être obtenus avec la NIH-EXAMINER, selon la théorie de réponse aux items (*item response theory*) utilisée par Kramer *et al.* (2014). Ces données ont été calculées grâce à un script informatisé programmé avec le logiciel R. Cette théorie prend en compte la variance accordée à chaque sous-test dans l'obtention du score de chaque facteur.

Afin de vérifier la comparabilité en termes d'âge de notre échantillon à celui de l'étude de validation originale (Kramer *et al.*, 2014), un test de comparaison de moyennes (test Z) a été réalisé. Pour vérifier la comparabilité entre les deux échantillons au niveau de la proportion de filles et de garçons, un test de Chi-carré avec correction de Yates a été effectué.

Étant donné la faible taille de l'échantillon, et pour contrôler l'influence potentielle de la normalité des distributions sur les résultats, des analyses non-paramétriques ont été réalisées. Le seuil de signification pour l'ensemble des analyses est de 0,05 (α critique = 0,05). Des indices de la cohérence interne des tests seront obtenus par méthode de bissection pair-impair. Puisque la plupart des tests inclus dans la NIH-EXAMINER sont volontairement conçus avec un niveau de difficulté croissant, la méthode de bissection pair-impair, qui compare deux parties d'un même test, apparaît plus pertinente que celle de l'alpha de Cronbach, qui compare toutes les bisections possibles du test.

Selon l'organisation des FE stipulée par Kramer *et al.* (2014) et par le modèle de Miyake *et al.* (2000), le facteur exécutif global devrait corrélérer avec les facteurs principaux (contrôle cognitif et mémoire de travail). Des analyses de corrélations de Spearman seront réalisées pour explorer les liens entre ces facteurs, ainsi qu'entre les facteurs et leurs sous-tests respectifs.

Pour être considérées valides (validité convergente) et comparables à la version originale, les mesures obtenues à la NIH-EXAMINER devront corrélérer avec les mesures présumées évaluer les mêmes construits. Des analyses de corrélations de Spearman seront alors réalisées entre les scores à la NIH-EXAMINER et aux échelles du *BRIEF*, à Interférence couleurs-mots, à Séquences de chiffres, à Empan spatial et à Fluidité graphique. Un indice de validité de critère concomitante sera obtenu en évaluant le lien entre l'âge et les scores aux tests de la NIH-EXAMINER, avec des corrélations de Spearman.

Résultats

Avant de présenter les résultats psychométriques, la prochaine section présente les analyses descriptives de l'échantillon, en termes d'âge et de genre, les données descriptives

des résultats obtenus aux questionnaires et tests cognitifs, incluant les variables contrôles ainsi que les résultats à la NIH-EXAMINER.

Description d'échantillonnage

En termes d'âge et de genre (voir Tableau I), l'échantillon est comparable à celui de la version originale de Kramer *et al.* (2014). L'âge moyen des 17 enfants ayant participé à l'étude est de 9,9 ans (ÉT : 1,11) et celui des 337 enfants de la version originale est de 10,16 ans (ÉT : 3,5). En termes d'âge, aucune différence significative n'est observée entre l'échantillon du présent projet et celui de Kramer *et al.* (2014) ($z = -0,907, p = 0,4306$). Au niveau du genre, la proportion de filles ayant participé à la présente étude n'est pas significativement différente de l'échantillon de l'étude de validation originale ($\chi^2 = 0,202, p = 0,6531$).

Tableau I. Caractéristiques sociodémographiques de l'échantillon

Variables	n	%
Genre		
Filles	10	58,8
Garçons	7	41,2
Revenu familial annuel brut		
45 000 – 80 000 \$	5	29,5
100 000 – 140 000 \$	7	41,2
140 000 – 190 000 \$	5	29,5
Moyenne générale au dernier bulletin		
70 – 75 %	5	29,4
75 – 84 %	4	23,6
85 % et +	6	35,3
Total	17	100

Variables contrôles

Le Tableau II présente les résultats aux tests et questionnaires servant à évaluer le fonctionnement intellectuel et les symptômes d'inattention et d'hyperactivité/impulsivité. Les scores standardisés (STD) du *WISC-IV* et les RC du *ARS-IV* sont rapportés afin de situer l'échantillon par rapport aux normes. Aucun participant n'atteint le seuil clinique aux deux échelles du *ARS-IV* (Inattention : 93^e RC. Hyperactivité/impulsivité : 98^e RC).

Tableau II. Scores obtenus à Vocabulaire, Matrices et ARS-IV

Outils d'évaluation	<i>N</i>	<i>M</i> (<i>ÉT</i>)	<i>ETM</i>	<i>Min-Max</i>
<i>WISC-IV</i> (STD)				
Vocabulaire	17	11,41 (3,08)	0,8	7-16
Matrices	17	11,53 (2,81)	0,68	6-17
<i>ARS-IV</i> (RC)				
Inattention	17	50 (25)	0,57	1-90+
Hyperactivité/impulsivité	17	25-50 (25)	0,51	1-90
Score total	17	50 (25)	0,52	1-90

Note. ETM = Erreur-type de mesure. *WISC-IV*: Wechsler Intelligence Scale for Children – 4th edition. *ARS-IV*: ADHD Rating Scale – 4th edition. STD = score standard. RC = rang centile.

Le Tableau III présente les résultats aux différentes mesures des FE utilisés. Les scores normalisés sont présentés afin de situer les résultats des participants par rapport aux normes. Tel que rapporté dans le Tableau III, les scores obtenus par les participants aux différentes mesures (*BRIEF*, Interférence couleurs-mots, Séquences de chiffres et Empan spatial) se situent dans les limites de la moyenne selon l'âge. Pour Fluidité graphique, ce sont les scores bruts qui sont présentés, puisque les normes disponibles, conçues auprès d'un échantillon de 20 enfants seulement, sont limitées au nombre de graphiques effectués (Strauss *et al.*, 2006).

Tableau III. Scores obtenus aux différentes mesures de FE

Outils d'évaluation	<i>N</i>	<i>M (ÉT)</i>	<i>ETM</i>	<i>Min-Max</i>
<i>BRIEF</i> (scores T)				
Inhibition	16	44,69 (4,76)	1,19	37-55
Flexibilité	16	47,19 (7,94)	1,99	36-62
Contrôle émotionnel	16	46,13 (7,94)	1,99	38-65
Initiation	16	51,63 (8,47)	2,12	36-71
MdT	16	51,81 (9,05)	2,26	36-68
Planification/organisation	16	48,00 (8,26)	2,07	39-67
Organisation du matériel	16	48,38 (7,41)	1,85	38-69
Monitoring	16	46,38 (8,56)	2,14	34-61
Régulation comportementale	16	45,38 (5,68)	1,42	36-57
Régulation cognitive	16	49,19 (8,21)	2,05	37-69
Score total	16	47,50 (7,1)	1,78	38-62
Interférence couleurs-mots (STD)				
Inhibition	17	11,82 (2,38)	0,58	6-15
Flexibilité	10	10,5 (2,59)	0,82	5-14
SC <i>WISC-IV</i> (STD)	16	10,44 (1,37)	0,34	8-13
ESP <i>WNV</i> (scores T)	17	51,35 (7,15)	1,73	37-61
Fluidité graphique (scores bruts)				
Graphiques	16	37,88 (12,95)	3,24	17-56
Répétitions	16	6,06 (3,24)	0,81	1-12
Stratégies	16	10,31 (7,48)	1,87	1-23

Note. ETM = Erreur-type de mesure. *BRIEF*: Behavior Rating Inventory of Executive Function. MdT = mémoire de travail. *WISC-IV* = Échelle d'intelligence de Wechsler pour enfants et adolescents – 4^e Édition. SC = Séquences de chiffres. ESP = Empan spatial. *WNV* = Échelle non verbale d'aptitudes de Wechsler Canadienne. STD = score standard. RC = rang centile.

Fidélité

Cohérence interne. La fidélité en cohérence interne obtenue est acceptable pour la plupart des tests de la version traduite de la NIH-EXAMINER. La fidélité en cohérence interne du Test non-structuré, du Test d'appariement, du *N-Back*, du *Flanker* et du *CPT* est satisfaisante (voir Tableau IV). Les coefficients de ces tests sont comparables ou supérieurs à la version originale. Les coefficients de fidélité de Fluidité catégorielle et Comptage de cercles sont cependant plus faibles qu'attendu et n'atteignent pas le seuil acceptable de 0,60 (Hogan, 2017). Au questionnaire de normes sociales et de l'échelle comportementale, les résultats des participants sont trop homogènes et montrent un effet plafond, ce qui rend impossible l'analyse de la cohérence interne. La cohérence interne du score d'autocritique est également impossible à analyser, puisqu'il s'agit d'un score unique (différence entre la performance perçue et performance réelle à Fluidité catégorielle).

Tableau IV. Scores obtenus à la NIH-EXAMINER et coefficients de fidélité (cohérence interne)

Mesures (scores bruts)	<i>N</i>	<i>M</i> (<i>ÉT</i>)	<i>ETM</i>	r_{xx}	<i>Min-Max</i>
Appariement	17	7,36 (0,48)	0,12	0,66	7-8
Comptage de cercles	17	15,59 (3,62)	0,88	0,4	8-22
<i>N-Back</i>	15	2,02 (0,9)	0,23	0,82	0,02-3,67
<i>Flanker</i>	17	8,62 (0,42)	0,1	0,68	8-9
<i>CPT</i>	17	92,88 (4,46)	0,25	0,94	82-97
FVC	17	25,18 (6,17)	1,5	0,43	14-37
TNS	17	16,41 (6,02)	1,46	0,86	4-26
Autocritique	17	5,63 (4,3)	1,04	–	0,68-14,4
Normes 1	17	0,75 (0,68)	0,17	–	0-2
Normes 2	17	3 (1,93)	0,48	–	0-6

Note. r_{xx} = coefficients de corrélation en cohérence interne obtenus par méthode de bissection pair-impair. FVC = fluidité verbale catégorielle. TNS = test non structuré. Normes 1 = score de sous-adhérence aux normes sociales. Normes 2 = score de sur-adhérence aux normes sociales.

Validité

Validité de construit. Le Tableau V présente les corrélations entre les différentes mesures obtenues à la NIH-EXAMINER correspondant aux FE du modèle de Miyake *et al.* (2000) et aux FE supplémentaires de Kramer *et al.* (2014). Le score total de fonctionnement exécutif de la NIH-EXAMINER corrèle modérément et significativement avec les scores au facteur de contrôle cognitif et au facteur de mémoire de travail, associés aux FE primaires de Miyake *et al.* (2000). Tel qu'attendu, les facteurs de contrôle cognitif et de mémoire de travail sont également corrélés entre eux, mais plus faiblement. De même, selon l'organisation de FE proposée par Kramer *et al.* (2014), le score total de fonctionnement exécutif corrèle de façon modérée et significative avec le score obtenu au test Fluidité catégorielle. Le score total obtenu à la NIH-EXAMINER ne corrèle ni avec le Test non structuré ni avec le questionnaire de normes sociales.

De plus, la cohérence entre les facteurs choisis par Kramer *et al.* (2014) et les sous-tests qui en font partie est satisfaisante. D'ailleurs, le facteur de contrôle cognitif est corrélé significativement et modérément avec le Test d'appariement ($r = 0,471$; $p = 0,008$) et le *Flanker* ($r = 0,618$; $p = 0,001$). Toutefois, il ne corrèle pas avec le *CPT* ($r = 0,255$; $p = 0,168$). Le facteur de mémoire de travail corrèle avec les tests qui évaluent ce construit, soit le Comptage de cercles ($r = 0,377$; $p = 0,044$) et le *N-Back* ($r = 0,824$; $p < 0,001$).

Tableau V. Intercorrélations des facteurs principaux de FE, de la TNS et du Questionnaire de normes sociales de la NIH-EXAMINER (validité de construit)

Mesures	1	2	3	4	5	6	7	8
1. Score total	–							
2. FaF	0,41*	–						
3. FaCC	0,65**	0,14	–					
4. FaMdT	0,52**	0,13	0,42*	–				
5. FVC	0,49**	0,88**	0,22	0,19	–			
6. TNS	-0,08	0,09	-0,22	0,11	0,08	–		
7. Normes 1	0,04	0,26	-0,06	-0,15	0,31	-0,12	–	
8. Normes 2	0,01	0,05	0,05	0,17	0,08	-0,08	-0,37	–

Note. * $p \leq 0,05$ ** $p < 0,01$. En gras : les corrélations attendues selon la version originale (voir Annexe 1). FaF = facteur de fluidité. FaCC = facteur de contrôle cognitif. FaMdT = facteur de mémoire de travail. FVC = fluidité verbale catégorielle. TNS = Test non structuré. Normes 1 = score de sous-adhérence aux normes sociales. Normes 2 = score de sur-adhérence aux normes sociales.

Validité convergente. Puisque seuls les scores bruts de la NIH-EXAMINER sont disponibles, afin d'assurer la comparabilité des scores en validité convergente, ce sont les scores bruts des outils qui ont été utilisés. Le Tableau VI présente les corrélations entre les mesures de la NIH-EXAMINER et les échelles du *BRIEF* censées mesurer les mêmes construits. Tel que rapporté dans le Tableau VI, le score total à la NIH-EXAMINER, les mesures de contrôle cognitif, de mémoire de travail, de planification et de respect des normes sociales ne sont pas corrélées aux échelles cliniques du *BRIEF* mesurant les mêmes construits, sauf pour le Test d'appariement et la sous-échelle de flexibilité du *BRIEF*, qui corréleront ensemble de façon modérée.

Tableau VI. Corrélations entre les scores à la NIH-EXAMINER et le BRIEF (validité convergente)

Mesures	Inh.	Flex.	MdT	Plan.	Émo.	RComp.	RCogn.	BRIEF
NIH Tot.	-0,2	0,31	0,11	-0,1	0,01	0,01	0,06	0,01
FaCC	-0,37	0,19	0,13	-0,1	-0,15	-0,21	0,04	-0,09
FaMdT	-0,1	0,33	0,21	0,08	0,09	0,182	0,28	0,27
Appariement	-0,54**	0,02	0,15	-0,2	-0,1	-0,2	-0,11	-0,11
Comptage de	0,2	0,31	0,19	0,39*	-0,12	0,09	0,47*	0,35
<i>N-Back</i>	-0,15	0,38	0,08	0,08	0,16	0,13	0,16	0,13
<i>Flanker</i>	-0,2	0,1	-0,03	-0,08	-0,01	-0,16	0,01	-0,11
<i>CPT</i>	0,05	0,26	-0,05	0,16	-0,26	-0,03	0,14	0,05
FVC	-0,08	0,07	0,01	-0,53**	0,13	0,06	-0,28	-0,2
TNS	0,05	0,12	-0,25	-0,15	-0,17	0,12	-0,13	-0,08
Auto	-0,15	-0,09	0,17	-0,19	-0,06	0,01	-0,01	0,03
Normes 1	0,01	-0,11	0,1	-0,29	-0,04	-0,04	-0,16	-0,14
Normes 2	0,28	0,15	0,17	0,02	0,21	0,18	0,18	0,16

Note. * = $p \leq 0,05$ ** = $p < 0,01$. En gras : les corrélations attendues selon la version originale (voir Annexe 1). NIH Tot. = score total à la NIH-EXAMINER. FaCC = facteur de contrôle cognitif. FaMdT = facteur de mémoire de travail. FVC = fluidité verbale catégorielle. TNS = Test non structuré. Auto = Test d'autocritique. Normes 1 = score de sous-adhérence aux normes sociales. Normes 2 = score de sur-adhérence aux normes sociales. Inh = sous-échelle d'inhibition. Flex = sous-échelle de flexibilité. MdT = sous-échelle de mémoire de travail. Plan. = sous-échelle de planification/organisation. Émo. = sous-échelle de contrôle émotionnel. RComp. = Index de régulation comportementale. RCogn. = Index de régulation cognitive. BRIEF Tot. = score total au BRIEF.

Le Tableau VII présente les corrélations entre les mesures obtenues à la NIH-EXAMINER, les différents tests cognitifs évaluant les mêmes construits et l'âge. Le score de contrôle cognitif et certains sous-tests mesurant ce construit corrélaient avec les scores de vitesse à Interférence couleurs-mots, évaluant le même construit. Au niveau des scores de précision, certaines relations significatives sont ressorties. Le score de contrôle cognitif de la NIH-EXAMINER corréla négativement et de façon modérée avec le nombre d'erreurs non corrigées à la condition flexibilité d'Interférence couleurs-mots ($r = -0,588$; $p = 0,036$). Le *Flanker* (inhibition) corréla avec le nombre d'erreurs non corrigées à la condition flexibilité du même test ($r = -0,515$; $p = 0,05$). Aucun score de la NIH-EXAMINER évaluant la mémoire

de travail et la fluidité ne corrèle avec les scores obtenus aux autres outils d'évaluation de ces construit.

Validité de critère concomitante. Tel qu'attendu, un effet significatif de l'âge est observé sur le score total de fonctionnement exécutif et le score au facteur de contrôle cognitif de la NIH-EXAMINER. Toutefois, contrairement à ce qui était attendu, aucun effet significatif de l'âge n'a été observé sur le score au facteur de mémoire de travail, au test Fluidité catégorielle et au Test non structuré. Il est toutefois possible d'observer un effet significatif sur le test *Flanker*.

Tableau VII. Corrélations entre la NIH-EXAMINER, les tests cognitifs et l'âge (validité convergente et de critère)

Mesures	Inh.	Flex.	SCOI	ESPOI	FG	Âge
NIH Tot.	-0,3	-0,47	0,33	0,16	-0,03	0,39*
FaCC	-0,45*	-0,51*	0,28	0,00	0,06	0,39*
FaMdT	-0,2	-0,1	0,37	0,15	-0,05	0,14
Appariement	-0,41*	-0,2	-0,25	0,02	0,15	0,12
Comptage de	0,07	0,00	0,22	-0,03	-0,34	0,19
<i>N-Back</i>	-0,35	-0,11	0,35	0,31	0,2	0,13
<i>Flanker</i>	-0,3	-0,65**	0,29	0,16	0,08	0,60*
<i>CPT</i>	-0,19	-0,25	0,21	0,09	-0,04	0,18
FVC	-0,2	-0,11	0,28	-0,01	0,04	0,21
TNS	0,03	0,16	0,32	0,19	0,09	-0,14

Note. * = $p \leq 0,05$ ** = $p < 0,01$. En gras : les corrélations attendues selon la version originale (voir Annexe 1). NIH Tot. = score total à la NIH-EXAMINER. FaCC = facteur de contrôle cognitif. FaMdT = facteur de mémoire de travail. FVC = fluidité verbale catégorielle. TNS = Test non structuré. Inh. = temps (secondes) à la condition d'inhibition du test Interférence couleurs-mots. Flex. = temps (secondes) à la condition de flexibilité du test Interférence couleurs-mots. SCOI = nombre de chiffres rappelés correctement en ordre inverse. ESPOI = nombre de blocs rappelés correctement en ordre inverse. FG = nombre de graphiques réalisés au test Fluidité graphique.

Discussion

La présente étude visait à traduire en français la batterie d'évaluation des FE NIH-EXAMINER selon la méthode de Vallerand (1989) et à valider cette version auprès de 17 enfants québécois neurotypiques d'âge scolaire (8 à 12 ans) afin de répondre au manque d'outils d'évaluation adaptés aux enfants québécois. Elle présente donc des résultats préliminaires de validation, et ses tableaux ne constituent pas des normes. Les qualités psychométriques de la version traduite, telles que la fidélité en cohérence interne, la validité de construit, la validité convergente et la validité de critère, ont été explorées.

Synthèse des résultats

Les résultats préliminaires montrent une cohérence interne satisfaisante pour la majorité des tests de la NIH-EXAMINER, sauf pour Comptage de cercles et Fluidité catégorielle. La validité de construit explorée entre les différents facteurs de la batterie est satisfaisante et cohérente avec l'organisation de FE proposée par Miyake *et al.* (2000) et Kramer *et al.* (2014). Par ailleurs, les résultats n'ont pas démontré de lien significatif entre le facteur global et les mesures de planification ou de respect des normes sociales de la NIH-EXAMINER.

Plusieurs tests psychométriques ont été utilisés pour explorer la validité convergente. Les résultats se sont montrés concluants quant aux mesures de contrôle cognitif de la NIH-EXAMINER et Interférence couleurs-mots, confirmant ainsi nos hypothèses. Toutefois, les résultats ne permettent pas de distinguer l'inhibition de la flexibilité, alors qu'il s'agit de deux sous-composantes distinctes selon Miyake *et al.* (2000). Ces résultats demeurent cohérents avec ceux de l'étude de validation originale. Cependant, aucun lien suffisamment significatif entre les mesures de mémoire de travail de la NIH-EXAMINER et les autres outils d'évaluation (Séquences de chiffres et Empan spatial) n'a pu être relevé. L'absence de

corrélation entre Fluidité catégorielle de la NIH-EXAMINER et Fluidité graphique infirme également nos hypothèses. Enfin, la majorité des analyses ne permettent pas de montrer de liens suffisamment significatifs entre la NIH-EXAMINER et les comportements quotidiens associés aux FE (*BRIEF*). Le Test d'appariement est le seul qui corrèle de façon attendue avec le *BRIEF*.

Pour la validité de critère, les résultats montrent un effet significatif de l'âge sur le score au facteur global de la NIH-EXAMINER, au facteur de contrôle cognitif et au *Flanker*. Or, les résultats ne décèlent pas d'effet d'âge sur le facteur de mémoire de travail, Fluidité catégorielle et le Test non structuré, et infirment nos hypothèses.

Les résultats en validité convergente pour les comportements associés aux FE (*BRIEF*) ne se sont pas montrés aussi concluants qu'espérés et ne permettent donc pas de dépister les difficultés comportementales associées aux FE.

Critique du cadre théorique et de la conceptualisation de la NIH-EXAMINER

Cette section vise à discuter de la pertinence du choix de cadre théorique de la part de Kramer *et al.* (2014), ainsi que de l'importance de l'aspect développemental dans l'évaluation d'enfants. Suivra une critique de la méthode de scores factoriels à variance partagée employée par Kramer *et al.* (2014). Le fonctionnement exécutif tel que décrit par Miyake *et al.* (2000) est bel et bien présent dans sa forme cristallisée à l'âge adulte. Chez l'enfant, il semblerait que l'émergence de composantes liées aux FE ait lieu dès l'âge de trois ans mais que la distinction entre les différentes sous-composantes de FE ne serait pas encore possible à cet âge (Wiebe *et al.*, 2011). D'ailleurs, une récente étude réalisée auprès de 1020 participants âgés de trois à 21 ans a montré que les composantes de FE, évaluées à la NIH TOOLBOX (Gershon *et al.*,

2010), une batterie cognitive similaire à la NIH-EXAMINER, étaient indifférenciées chez les trois à six ans (Akshoomoff, Brown, Bakeman et Hagler Jr, 2018).

Les FE se différencieraient à partir de l'âge de six ans et leur organisation en trois composantes distinctes stipulée par Miyake *et al.* (2000) pourrait donc s'avérer juste pour décrire le fonctionnement exécutif durant l'enfance. Rappelons effectivement que l'étude de Lehto et ses collègues (2003) soutient l'utilisation d'un modèle à trois facteurs pour l'évaluation des FE chez les enfants.

Par ailleurs, Wasserman et Wasserman (2013) rappellent l'importance d'établir des trajectoires développementales pour l'évaluation des processus de FE durant l'enfance. En effet, les différentes composantes de FE suivent des courbes de développement distinctes. Selon Anderson (2002), l'attention surviendrait assez tôt dans l'enfance, tandis que la flexibilité, la sélection d'un but (*goal setting*) et le traitement de l'information progresseraient de 7 à 9 ans pour atteindre leur maturité vers 12 ans. L'intégration plus étoffée du contrôle cognitif à partir des sous-composantes se ferait surtout à l'adolescence (Anderson, 2002). Ainsi, certaines FE pourraient être moins bien développées que d'autres et moins intégrées en fonction de l'âge, sans que cela ne témoigne d'un déficit franc des FE. Il apparaît pertinent de connaître l'évolution développementale de chaque sous-composante de FE pour l'évaluation. À l'âge adulte, il est conseillé que l'évaluation des FE prenne en compte l'aspect unitaire (facteur global de fonctionnement exécutif) tout autant que l'aspect fragmenté (sous-composantes de FE) (Miyake *et al.*, 2000), tandis qu'à l'enfance, des trajectoires développementales particulières et une intégration différente des sous-composantes pourraient remettre en cause cette approche d'évaluation (Wasserman et Wasserman, 2013). Effectivement, en raison de différences neuroanatomiques, il se pourrait que l'interdépendance

des sous-composantes soit moins bien établie à l'enfance et qu'elle diffère même selon l'âge de l'enfant.

Ainsi, il est possible de déduire que la dissociation entre les composantes du modèle théorique utilisé par la NIH-EXAMINER se fait de manière graduelle entre l'âge de 6 et 8 ans, approximativement. Il est crucial de bien garder cette période de transition en tête lors de l'évaluation des FE d'enfants. Bien que les enfants de notre échantillon aient au moins 8 ans, il est possible de croire que la dissociation progressive des composantes de FE ait eu un impact sur nos résultats. De futures études pourraient d'ailleurs répliquer l'étude de Lehto *et al.* (2003) en utilisant plutôt les scores factoriels de la NIH-EXAMINER obtenus selon différentes tranches d'âge. Cela permettrait d'établir des hypothèses sur les trajectoires développementales des FE.

Puisque les enfants plus jeunes ont des FE non-dissociées, il est avisé de considérer l'âge de l'enfant pour le choix d'un modèle de FE à un facteur (score global de FE) ou d'un modèle à trois facteurs (contrôle cognitif, mémoire de travail et flexibilité). D'ailleurs, l'utilisation des modèles factoriels confirmatoires, tels qu'utilisés par Miyake *et al.* (2000) et Lehto *et al.* (2003), et implémentés dans la NIH-EXAMINER, comprend plusieurs avantages, dont celui d'atténuer le problème d'impureté des tâches (Packwood, Hodgetts et Tremblay, 2011), c.-à-d. le fait qu'une tâche mesure plusieurs fonctions différentes en même temps, ce qui contamine le score total obtenu à une tâche avec d'autres fonctions cognitives que celles d'intérêt. Il est donc crucial d'avoir plusieurs tâches qui mesurent la même fonction lorsqu'on évalue les FE. En se référant à la théorie de réponse par item pour obtenir des scores factoriels, la variance associée à chacune des tâches impliquées dans l'obtention du score total de chacun des facteurs de la batterie est prise en compte et amoindrit le problème d'impureté. En effet,

l'obtention de scores factoriels diminuent l'influence de variables confondantes, comme d'autres fonctions cognitives impliquées dans la tâche (p.ex. le langage) ou des caractéristiques techniques liées à la nature de la tâche (Miyake et Friedman, 2012; Yong et Pearce, 2013).

Malgré ces avantages, la méthode d'analyses factorielles confirmatoires utilisée par Miyake *et al.* (2003) et par Lehto *et al.* (2003) peut poser problème. D'ailleurs, cette méthode à variance partagée se heurte à plusieurs critiques, car elle pourrait masquer certaines distinctions plus fines mais particulièrement importantes pour l'évaluation entre les facteurs obtenus (Delis, Jacobson, Bondi, Hamilton et Salmon, 2003; Larrabee, 2003). La méthode à variance partagée serait inappropriée pour des populations qui ne présentent aucune difficulté, car un effet plafond peut survenir, ou encore pour des participants présentant des difficultés cognitives sévères, puisqu'un effet plancher est alors probable. Elle risque alors de fournir des résultats non-significatifs au plan statistique empêchant la variance nécessaire à la distinction des facteurs, en raison de la trop grande homogénéité des scores. Il ne sera pas possible de déceler les nuances entre les scores aux sous-tests propres aux différentes fonctions cognitives.

De plus, l'analyse factorielle confirmatoire requiert l'identification de facteurs prédéterminés, ce qui ne permet pas de rendre compte de toute l'étendue et de la complexité du construit des FE. En forçant un modèle à trois composantes comme celui de Miyake *et al.* (2000), il est possible que certaines nuances, voire certaines fonctions, ne soient pas capturées dans les scores factoriels. Par exemple, la planification n'est pas incluse dans les composantes principales du modèle de Miyake *et al.* (2000) ni dans la conceptualisation de Lehto *et al.* (2003). Pourtant, elle serait le seul facteur commun de FE à l'enfance selon certains auteurs (Anderson, 2002) et fait partie de bon nombre de modèles conceptuels des FE (Burgess et

Shallice, 1997; Gioia et Isquith, 2004; Goldstein et Naglieri, 2013; Lezak, 1995; Shallice et Burgess, 1996; Stuss et Benson, 1986; Zelazo *et al.*, 1997). Kramer *et al.* (2014) ont justement inclus une mesure de planification du fait qu'elle est souvent associée aux FE dans la littérature. Comme les scores de la tâche de planification ne sont pas prépondérants dans le calcul des scores factoriels de la NIH-EXAMINER, il est pertinent de se demander si un modèle non-contraint par des a priori théoriques comme ceux utilisés dans le développement de la NIH-EXAMINER serait plus adéquat.

De plus, il est de mise de questionner le choix et la représentativité des tâches associées à chacune des FE du modèle de Miyake *et al.* (2000). En effet, bien que le cadre conceptuel de la NIH-EXAMINER prenne en compte l'inhibition, il semblerait qu'il existe plus de huit formes d'inhibition distinctes selon le type de pathologie associée (Nigg, 2000). L'inhibition évaluée par la NIH-EXAMINER est loin d'être aussi étoffée, car elle réfère simplement à l'inhibition comportementale, soit la suppression d'une réponse automatique au profit d'une réponse contrôlée.

De même, l'implication dans les FE de la sous-composante de mémoire de travail apparaît beaucoup plus importante que la simple capacité à traiter et manipuler l'information à court terme. Effectivement, la composante exécutive de mémoire de travail, le Centre exécutif, est conçu comme un gestionnaire de ressources attentionnelles divisées en sous-composantes et orientées vers un but (Baddeley, 1996, 2012). De plus, la mémoire de travail comprend la capacité de mise à jour continue de l'information à court terme pour s'y ajuster (Miyake *et al.*, 2000). Toutefois, les sous-tests de la NIH-EXAMINER qui l'évaluent ne font référence qu'au traitement et au maintien de l'information à court terme (Comptage de cercles et *N-Back*). D'ailleurs, le test de Comptage de cercles où l'enfant doit rappeler dans l'ordre les

cercles comptés, est en fait une tâche d'empan direct qui évalue la rétention en mémoire à court-terme sans aucune manipulation de l'information. Il appert que toute la complexité de la mémoire de travail n'est alors pas prise en compte par les deux sous-tests de la batterie.

Critique de l'étude de validation originale de Kramer *et al.* (2014)

Cette section présentera certaines failles notées dans l'étude de validation originale ainsi que les limites associées à certains sous-tests, soit le questionnaire de normes sociales, l'échelle d'observation comportementale et la tâche d'*insight* (autocritique).

Les concepteurs de la NIH-EXAMINER ont inclus d'autres mesures de FE que celles recommandées par l'étude de Miyake *et al.* (2000): la planification, la fluidité verbale, l'autocritique (*insight*), le respect des normes sociales et le contrôle comportemental. Ceci permet de prendre en compte d'autres FE souvent incluses dans l'évaluation neuropsychologique. Toutefois, plusieurs données psychométriques de la batterie pour enfants ne sont disponibles ni dans l'article de validation, ni dans le manuel de l'examineur.

De plus, le questionnaire de normes sociales a seulement été validé auprès de groupes de personnes âgées neurotypiques et de groupes présentant une démence de type fronto-temporale. Deux items ont été retirés du questionnaire dans la version enfants sans que les raisons soient explicitées. Lors de la présente étude, un effet plafond est obtenu à ce questionnaire (tous les enfants répondent de façon homogène sans aucune erreur). Une étude de validation de la version traduite en français de la NIH-EXAMINER réalisée auprès d'un échantillon d'adultes neurotypiques montre le même effet de plafonnement à cette tâche (Dumont, 2017). Les comportements sociaux inappropriés rapportés dans le questionnaire sont très facilement identifiables pour des personnes neurotypiques. Celui-ci n'apparaît pas suffisamment discriminant pour déceler des atteintes plus fines et est fortement entaché par la

désirabilité sociale. Une tâche comme celle de Théorie de l'esprit de la NEPSY-II ou de raisonnement moral (Beauchamp et Dooley, 2013), qui mesure plus finement des aspects de la cognition sociale, pourrait le remplacer.

Quant à l'échelle comportementale, un effet plafond (homogénéité des scores) est aussi observé pour l'échantillon de la présente étude et pour celui de Dumont (2017). Les comportements visés par cette échelle sont observés pendant la passation de la batterie, alors que les enfants sont en situation un à un avec l'évaluateur pendant moins d'une heure. Rappelons qu'il s'agit d'enfants ayant de bons résultats scolaires et pour lesquels aucune difficulté à l'école n'est rapportée. Il est possible de présumer que peu de comportements problématiques se manifesteront durant la passation de la NIH-EXAMINER (45 minutes).

La tâche d'*insight* (autocritique) intégrée à la NIH-EXAMINER semble également poser problème. D'abord, aucune instruction n'est donnée pour obtenir un score quantifiable. Dans la présente étude, la différence absolue entre la performance perçue de l'enfant à Fluidité et sa performance réelle a été calculée pour obtenir le score. Cette méthode est problématique, car elle nécessiterait d'avoir des données normatives de la performance des enfants à la tâche de Fluidité, ce qui n'est ni notre cas ni celui de l'étude de validation originale de Kramer *et al.* (2014). De plus, la tâche d'autocritique réfère à l'autoévaluation de l'enfant à une seule tâche. Cette autoévaluation montre plutôt la capacité de l'enfant à se comparer à ses pairs. Aussi, Kramer *et al.* (2014) estiment que cette mesure évalue la régulation comportementale de l'enfant. Le fait qu'elle soit liée à l'échelle de flexibilité du BRIEF (Krueger *et al.*, 2011) plutôt qu'à l'échelle de régulation comportementale nous semble arbitraire. En plus, cette tâche ne fait aucunement référence à la capacité d'introspection, de connaissance de soi et de perception juste de ses limites ou de ses forces qui font hypothétiquement partie intégrante de

la notion d'*insight* (Markova et Berrios, 1992). Des modifications à la tâche originale ont été réalisées dans l'étude de validation auprès de participants adultes (Dumont, 2017), dans le but de minimiser l'effet de comparabilité sociale. Les participants devaient estimer le nombre d'erreurs commises à une des tâches disponibles dans la version pour adultes (Anti-saccades). Malgré cette modification, Dumont (2017) n'a pas obtenu les résultats attendus. En raison des lacunes susmentionnées, il nous paraît nécessaire de reconsidérer la pertinence de cette tâche.

Limites de la présente étude

Méthodes de traduction. La présente étude comporte certaines limites méthodologiques. La traduction selon la méthode de Vallerand (1989) est plus optimale qu'une simple traduction directe de la langue originale à la langue cible. Toutefois, certains standards, et plus particulièrement l'adaptation culturelle des outils d'évaluation, auraient pu être employés. La culture a un effet démontré sur la façon d'aborder les tâches, sur la façon de réagir à la situation-même d'évaluation neuropsychologique ainsi que sur la validité de contenu des tâches (Ardila, 2005; Kennepohl, 1999). En raison de l'importance de cet impact, la traduction ne devrait être que la première étape du processus d'adaptation d'un outil (Beaton, Bombardier, Guillemin et Ferraz, 2000). D'ailleurs, Beaton *et al.* (2000) recommandent un processus d'adaptation culturelle en cinq étapes : (1) traduction directe, de l'anglais au français, (2) la synthèse de la première traduction, (3) la traduction inversée, (4) l'évaluation par un comité d'experts et (5) le pré-test. Un rapport explicatif devrait être produit à chacune des étapes de ce processus pour documenter chacune des décisions prises. De plus, lors de la traduction directe (1), il est préférable d'avoir recours à deux traducteurs bilingues dont la langue maternelle est le français et que, l'un d'eux ne soit pas familier avec le construit mesuré par l'outil. Par la suite, une synthèse est réalisée en présence des deux traducteurs et

d'un observateur externe afin d'obtenir un consensus sur les items posant problème (2). Une traduction inversée est alors réalisée (3), du français vers l'anglais, par deux traducteurs différents et naïfs dont la langue maternelle est l'anglais. Puis un comité d'experts d'environ 30 personnes, composé de méthodologistes, de professionnels de la santé, de linguistes et des traducteurs impliqués, devrait se réunir pour le consensus d'une version pré-finale (4). Un pré-test chez 30 à 40 personnes serait réalisé afin d'évaluer cette version (5). Lorsque le processus est terminé, un rapport est remis aux développeurs qui jugent de la rigueur du processus d'adaptation et non de la qualité de la traduction. C'est seulement à la suite de ces cinq étapes que le processus de validation peut débuter (Anthoine, Moret, Regnault, Sébille et Hardouin, 2014). Il est clair que ce processus n'a pas été respecté dans cette étude. Le manque de ressources financières et humaines explique en partie cette faille.

Échantillonnage. La faible taille de l'échantillon final ($n = 17$) est une limite importante de l'étude, car le poids statistique de chaque participant de l'échantillon est plus élevé. Un participant qui obtient un score atypique affecte de façon plus marquée la distribution d'un échantillon restreint. Dans le cas où les participants sont similaires, le peu de variabilité entre eux empêche de détecter de plus petits liens qui pourraient facilement l'être avec un échantillon plus grand. Si le but de l'étude était d'obtenir des corrélations significatives ($p < 0,05$), avec une puissance de 0,8 et un coefficient de corrélation d'environ 0,35, l'échantillon visé serait de 59 à 62 enfants (Faul, Erdfelder, Buchner et Lang, 2009). Cependant, lors d'un processus de validation, le but n'est pas d'obtenir des résultats significatifs au point de vue statistique, mais suffisamment robustes pour pouvoir affirmer que l'outil mesure bel et bien ce qu'il doit mesurer. Les recommandations sont alors plus strictes puisque l'outil pourrait être utilisé dans la population. La structure factorielle du construit

mesuré par l'outil permet de rendre les résultats plus fiables. La taille minimale d'échantillon requise pour les analyses factorielles est de 150 à 200 personnes, alors que 300 personnes et plus semblent convenables pour effectuer les statistiques requises (Anthoine *et al.*, 2014; Arafat, Chowdhury, Qusar et Hafez, 2016; MacCallum, Widaman, Zhang et Hong, 1999). Puisque la NIH-EXAMINER possède un faible nombre de facteurs et un faible nombre de sous-tests par facteurs, il est préférable de viser une taille d'échantillon d'environ 300 personnes lors de sa validation (MacCallum *et al.*, 1999). L'échantillon de l'étude originale respecte cette norme avec 337 enfants. Dans la présente étude, les données de 17 sujets ont été retenues, ce qui est très insuffisant pour valider la structure factorielle d'un outil.

La faible variabilité entre les participants est une autre limite importante. En effet, puisque les participants de cette étude sont des enfants neurotypiques pour la plupart issus de milieux socio-économiques aisés, ils ne représentent pas la variabilité interindividuelle réelle de la population québécoise. Dans une certaine mesure, ceci pourrait expliquer l'effet plafond obtenu au questionnaire de normes sociales, à l'échelle comportementale et au *CPT*.

Malgré les lacunes quant à la traduction de l'outil, ces données peuvent s'inscrire dans un pré-test, soit l'étape cinq du processus d'adaptation culturelle selon Beaton *et al.* (2000). Le pré-test devrait être poursuivi avec un minimum de 13 sujets supplémentaires (30 participants au total) avant que le processus de validation ne soit véritablement entamé. Les résultats obtenus dans cette étude demeurent pertinents, mais ne sont que préliminaires et s'inscrivent dans un devis exploratoire.

Choix de tests pour la validation. Une autre limite méthodologique concerne le choix des outils pour la validation convergente de la version traduite de la NIH-EXAMINER. D'abord, l'Empan spatial a été administré à l'ordinateur, alors que cette version du test n'a

jamais été validée. Le *5-point Test* a également été administré sur tablette, alors qu'il s'agit d'une modification de l'original, dont le processus de validation n'est pas encore complété. Idéalement, les versions papiers-crayons validées auprès d'enfants neurotypiques devraient être utilisées pour explorer la validité convergente.

De plus, le *BRIEF* ne semble pas la mesure la plus adaptée à l'évaluation des FE, puisqu'il mesure plutôt les comportements au quotidien. En ce sens, Barkley et Murphy (2010) ont montré que les résultats aux tests neuropsychologiques de FE ne permettent pas de prédire les difficultés fonctionnelles chez des personnes TDA/H. Les tests utilisés ne sont pas suffisamment valides écologiquement, comparativement aux mesures auto-rapportées. Ils mesurent les processus cognitifs de FE, alors que les questionnaires comme le *BRIEF* permettent d'évaluer l'impact comportemental des FE dans le quotidien. Une recension des écrits de 2013 a montré que les relations entre le *BRIEF* et les tests psychométriques de FE étaient pauvres, avec une faible proportion de corrélations significatives (19% des études) ou des coefficients de corrélation moyens de 0,15 (Toplak, West et Stanovich, 2013). Ainsi, les résultats obtenus dans la présente étude entre la NIH-EXAMINER et le *BRIEF* pourraient être expliqués par le fait que ces mesures ne sont pas conçues pour évaluer le même construit. Pour évaluer la planification, une tâche standardisée, telle que les labyrinthes de la WISC-III ou les 6 éléments de la BADS-C, aurait pu être incluse dans le protocole de recherche pour explorer la validité convergente plutôt que la sous-échelle de planification au *BRIEF*. Pour les autres mesures de FE, le recours à une batterie validée et normée auprès du même échantillon aurait pu être intéressante à utiliser. En ce qui a trait à la validité divergente, des mesures supplémentaires auraient pu être ajoutées au protocole de recherche comme le *CVLT-II* (Delis, Kramer, Kaplan et Ober, 2000) évaluant la mémoire ou le *BNT* (Kaplan, Goodglass et

Weintraub, 1983), évaluant le langage expressif.

La NIH-EXAMINER présente certains atouts pour la recherche clinique et l'évaluation neuropsychologique, pour obtenir rapidement des mesures de différentes FE. Elle fournit un vaste éventail de scores différents pour chacun des sous-tests (taux de réponses, erreurs commises, vitesse, etc.), se fonde sur un modèle empiriquement validé (Lehto *et al.*, 2003; Miyake *et al.*, 2000) et permet de dresser un profil complet de plusieurs FE en moins d'une heure. De plus, elle est gratuite, libre d'accès et facile d'utilisation quant à la passation et la cotation.

Depuis sa conceptualisation en 2014, la version originale de la NIH-EXAMINER a été utilisée auprès de groupes cliniques. Ces études, bien que sommaires, ont montré des évidences neuroanatomiques qui soutiennent la pertinence d'un tel outil pour la recherche clinique. En effet, les facteurs de FE mesurés à la NIH-EXAMINER se sont montrés sensibles à la présence d'atteintes cérébrales au niveau du corps calleux parmi un groupe de 32 enfants souffrant d'anémie falciforme (Schatz, Stancil, Katz et Sanchez, 2014). Le facteur de fonctionnement exécutif global obtenu à la NIH-EXAMINER prédit les comportements associés aux FE, mesurés par le *Frontal System Behavioral Scale (FrsBE)*, (Grace et Malloy, 2001), et un faible score à la NIH-EXAMINER est également lié à des atteintes frontales (Possin, LaMarre, Wood, Mungas et Kramer, 2014). Une étude auprès d'un groupe de 32 enfants TDA/H a permis de montrer des scores significativement plus faibles au facteur de mémoire de travail de la NIH-EXAMINER comparativement à un groupe contrôle.

À notre connaissance, aucune autre étude de validation de cette batterie chez une population d'enfants québécois n'a été réalisée. Cependant, une thèse récemment publiée avait pour objectif de valider la NIH-EXAMINER traduite auprès d'adultes et de la rendre

disponible à des fins de recherche (Dumont, 2017). Les résultats obtenus par 60 participants neurotypiques bilingues aux trois versions disponibles de la NIH-EXAMINER respectent la structure factorielle du construit postulé, sauf pour Fluidité verbale. Selon Dumont, ces résultats à Fluidité s'expliquent par le niveau d'aisance en anglais par rapport au français, langue maternelle des participants.

Études futures

Des études futures permettraient de consolider les résultats concernant les qualités psychométriques de l'outil. Afin d'explorer davantage la fidélité, il pourrait être intéressant de vérifier l'équivalence entre les différentes versions disponibles de la NIH-EXAMINER (A, B et C), ainsi que la stabilité de la batterie à différents temps de mesure. Pour obtenir plus d'informations sur la validité de critère concomitante, il serait utile de soumettre cette batterie à des groupes cliniques d'enfants présentant des atteintes exécutives (p.ex. TDA/H, trouble souvent associé à un déficit des FE selon Barkley (1997) et Sonuga-Barke (2003, 2005), dyslexie, TSA, SGT, etc.).

Conclusion

L'objectif de cet essai visait à traduire et à explorer les qualités psychométriques de la batterie d'évaluation des FE, la NIH-EXAMINER. Dix-sept participants (10 filles et 7 garçons) constituent l'échantillon de cette étude exploratoire de validation. La batterie NIH-EXAMINER, d'autres tests évaluant les FE et des tests de fonctionnement intellectuel leur ont été administrés. Des données concernant l'impact des FE dans le quotidien et les symptômes d'inattention et d'hyperactivité/impulsivité liés au TDA/H ont également été recueillies. Les résultats d'analyses non-paramétriques ont montré une fidélité en cohérence interne

satisfaisante, une bonne validité de construit pour la majorité des tests et un effet d'âge similaire à celui de la version originale. Toutefois, certains résultats ne sont que partiellement satisfaisants pour la validité convergente. Le processus de validation de la NIH-EXAMINER bénéficierait d'être poursuivi auprès de plus grands échantillons et de groupes cliniques. La NIH-EXAMINER présente plusieurs avantages, dont l'informatisation de la majorité des tâches, ce qui diminue les erreurs de passation et de cotation, le libre-accès à son matériel, l'obtention de scores factoriels et l'appui sur un modèle conceptuel validé empiriquement. L'adaptation d'un tel outil à la population francophone du Québec est hautement pertinente pour la tenue de recherches scientifiques au sujet des FE et, au niveau clinique, pour l'évaluation de difficultés exécutives.

**Annexe 1 : Description de la NIH-EXAMINER et données psychométriques de la version
originale**

Facteur global et facteurs principaux

La fidélité test-retest du score total de fonctionnement exécutif chez les moins de 18 ans est modérément élevée, soit de 0,87 après quatre mois et de 0,80 après un an. En termes de validité convergente, les données d'enfants n'étant pas disponibles dans le manuel de l'examineur, les données obtenues auprès d'adultes neurotypiques ($n = 74$) permettent d'observer une corrélation positive significative entre le score total de fonctionnement exécutif et le test Interférence couleurs-mots (*Color-Word Interference Test* ; *Delis-Kaplan Executive Function System (DKEFS)* ; Delis, Kaplan et Kramer, 2001). De plus, le score total de fonctionnement exécutif obtenu corrèle significativement avec le score total obtenu à l'Inventaire d'évaluation comportementale des fonctions exécutives (*Behavior Rating Inventory of Executive Function (BRIEF)* ; Gioia, Isquith, Guy et Kenworthy, 2000) par des enfants neurotypiques ($n = 404$). En termes de validité de critère concomitante, le score total de fonctionnement exécutif obtenu à la NIH-EXAMINER permettrait de distinguer les enfants neurotypiques des enfants des groupes cliniques, le score total de fonctionnement exécutif obtenu par le groupe contrôle ($n = 310$) étant significativement meilleur que celui obtenu par les enfants des groupes cliniques ($n = 139$).

Pour le score de contrôle cognitif, regroupant le Test d'appariement, le *Flanker* et *CPT*, chez les moins de 18 ans, la fidélité test-retest est modérément élevée, avec un coefficient de corrélation de 0,83 après 4 mois et de 0,77 après un an. De plus, le score de contrôle cognitif corrèle significativement avec le score total du *BRIEF* et les sous-échelles de flexibilité et d'inhibition. Le score de contrôle cognitif corrèle également avec le test d'Interférence couleurs-mots chez l'adulte. Au score de contrôle cognitif, comprenant les tests évaluant l'inhibition et la flexibilité cognitive, les enfants du groupe contrôle ($n = 368$)

obtiennent une performance significativement meilleure que les enfants des groupes cliniques ($n = 119$). Un effet significatif de l'âge est observé sur le score total de fonctionnement exécutif chez 278 enfants neurotypiques, et au score de contrôle cognitif chez 368 enfants neurotypiques. L'âge a donc été considéré comme covariable par Kramer *et al.* (2014) dans le calcul de certains indices psychométriques de fidélité et de validité.

Au facteur de mémoire de travail, regroupant les tests Comptage de cercles et *N-Back*, la fidélité test-retest est modérément élevée à faible, avec des coefficients de corrélation de 0,77 après 4 mois et 0,52 après un an. Le score au facteur de mémoire de travail corrèle positivement et significativement avec le test Séquences de chiffres en ordre inverse (*Digit Span Backward* ; *Wechsler Adult Intelligence Scale – 4th edition (WAIS-IV)* (Wechsler, 2008). Le score de mémoire de travail corrèle également avec la sous-échelle de mémoire de travail et le score total du *BRIEF*, chez des enfants neurotypiques ($n = 341$). En termes de validité divergente, le facteur de mémoire de travail n'est pas corrélé à des outils qui évaluent d'autres fonctions, tels que le *Boston Naming Test (BNT)* (Kaplan *et al.*, 1983), évaluant le vocabulaire expressif, et le *California Verbal Learning Test – second edition (CVLT-II)* (Delis *et al.*, 2000), évaluant la mémoire épisodique verbale. Le score obtenu au facteur mémoire de travail permet de distinguer les enfants des différents groupes cliniques ($n = 119$), qui obtiennent un score significativement plus faible que ceux du groupe contrôle ($n = 368$). De plus, le score obtenu au facteur mémoire de travail par les enfants du groupe clinique ayant un TDA/H ($n = 32$) est significativement plus faible que celui des enfants neurotypiques ($n = 164$). Un effet de l'âge significatif est également observé sur le score obtenu en mémoire de travail par des enfants neurotypiques de moins de 15 ans ($n = 341$).

Flexibilité cognitive

Test d'appariement. Ce test évalue la capacité de l'enfant à s'adapter aux différentes demandes de l'environnement et à ajuster sa réponse. L'enfant doit appairer des stimuli qui lui sont présentés à l'écran selon la forme ou la couleur. La consigne (par ex. forme ou couleur) lui est fournie par l'enregistrement sonore. Tout au long du test, la consigne change. Il y a trois conditions au test. Dans la première condition, l'enfant apparie les stimuli selon la couleur seulement. Dans la deuxième, il les apparie selon la forme seulement et dans la troisième, il les apparie en alternant entre la forme et la couleur. Pour chaque condition, des mesures d'exactitude de la réponse et de temps de réaction sont obtenues. Ce test possède une fidélité en cohérence interne élevée, avec des alphas de Cronbach entre 0,86 et 0,97, dépendamment de la condition administrée.

Mémoire de travail

Comptage de cercles. Ce test consiste à compter les cercles bleus qui apparaissent à l'écran parmi des éléments de distraction, des carrés bleus et des cercles verts. Ensuite, des configurations différentes de cercles et d'éléments de distraction apparaissent à l'écran. L'enfant doit maintenir en mémoire le nombre de cercles bleus comptés et rappeler ensuite les nombres de cercles comptés dans l'ordre, à la fin de l'essai. Ce test comprend six essais et le nombre de configurations de cercles bleus et de distractions augmente à chaque essai. Le premier essai comprend deux configurations différentes à rapporter et le dernier essai en comprend sept. Le nombre total de bonnes réponses données est noté. Le test de Comptage de cercles possède une fidélité en cohérence interne acceptable, avec un alpha de Cronbach de 0,69.

N-Back. Ce test consiste à comparer la localisation d'une forme présentée à une autre forme présentée juste avant sur l'écran d'ordinateur. Un carré blanc est présenté à l'enfant

parmi des éléments de distraction. Ensuite, le même carré blanc est présenté à nouveau parmi d'autres éléments de distraction. L'enfant doit alors appuyer sur la flèche de gauche si ce carré est au même emplacement que le carré précédent et sur la flèche de droite sinon. Le test comprend 30 essais, dont 20 essais où les carrés ne sont pas au même emplacement et 10 essais où ils le sont. Le nombre total de bonnes réponses données est noté. Il n'y a cependant pas de données disponibles concernant les qualités psychométriques du *N-Back* dans le manuel de l'examineur fourni par les concepteurs.

Inhibition

Flanker. Le *Flanker* consiste à présenter cinq poissons sur un écran d'ordinateur. L'enfant doit se concentrer sur le poisson du milieu qui nage dans une certaine direction. La direction dans laquelle le poisson nage peut être identique ou différente de celle des autres poissons. L'enfant doit appuyer sur la flèche de droite si le poisson du milieu nage vers la droite et sur la flèche de gauche s'il nage vers la gauche, sans se laisser influencer par la direction de nage des quatre poissons environnants. Des mesures d'exactitude (nombre de bonnes réponses) et de vitesse (temps de réponse) sont obtenues. Le *Flanker* possède une fidélité en cohérence interne modérément élevée à élevée, avec des alphas de Cronbach entre 0,80 et 0,97, dépendamment de la condition (par ex. congruente ou incongruente).

Continuous Performance Test (CPT). Le *CPT* consiste à répondre à un stimulus particulier (cible) et à inhiber une réponse en présence d'autres stimuli. Dans le présent test, l'enfant doit appuyer sur la barre d'espacement lorsqu'une étoile cible est présentée à l'écran d'ordinateur, mais ne pas appuyer si une autre forme que l'étoile est présentée. Dans 80% des cas, l'image présentée est l'étoile à cinq branches (cible). Dans 20% des cas, ce sont des images semblables mais différentes qui sont présentées à l'enfant (par ex. l'étoile à quatre

branches). Des mesures de commissions (sélection de non-cibles) sont obtenues. La fidélité en cohérence interne du *CPT* est acceptable, avec un alpha de Cronbach de 0,64.

Concernant la validité convergente, les tests évaluant l'inhibition de la NIH-EXAMINER corrént positivement et significativement avec le test Interférence couleurs-mots chez 75 participants adultes. En termes de validité divergente, les tests évaluant l'inhibition ne sont pas corrélés à des outils qui évaluent d'autres fonctions, tels que le *BNT* et le *CVLT-II*.

Fluidité verbale

Fluidité catégorielle. La fluidité verbale s'évalue en demandant à l'enfant de générer le plus de mots possible en une minute et dans une catégorie donnée (par ex. animaux ou légumes). Des mesures des bonnes réponses données, des répétitions et des violations de règles (par ex. mot d'une autre catégorie que celle demandée) sont notées. Ce test possède une fidélité en cohérence interne et un accord inter-juges modérément élevés avec un alpha de Cronbach et un coefficient de corrélation de 0,78. De plus, la fidélité test-retest du score de fluidité est modérément élevée à faible avec des coefficients de corrélation de 0,76 après quatre mois et de 0,73 après un an pour les moins de 18 ans. Ce test corréle fortement avec le test Fluidité graphique (*Design fluency, DKEFS*) et modérément avec les tests Séquences de chiffres en ordre inverse et Interférence couleurs-mots chez un groupe de participants adultes plus âgés ($n = 302$) présentant ou non une atteinte cognitive (par ex. trouble cognitif léger, aphasie progressive, démence de type Alzheimer). Cela suggère donc que le test Fluidité catégorielle possède une bonne validité convergente. Toutefois, le score de fluidité verbale catégorielle obtenu à la NIH ne corréle pas de façon significative avec le score total obtenu au *BRIEF* chez 398 enfants. En termes de validité divergente, les tests évaluant la fluidité ne sont

pas corrélés à des outils qui évaluent d'autres fonctions, tels que le *BNT* et le *CVLT-II*. Un effet significatif de l'âge est également observé sur le score obtenu au test Fluidité catégorielle par des enfants neurotypiques ($n = 277$).

Planification

Test non-structuré. Dans ce test, trois cahiers de quatre exercices chacun à compléter en trois minutes sont remis à l'enfant. Les exercices suggérés sont des labyrinthes, des cibles à biffer, des dessins à retracer et des symboles à recopier. Certains exercices prennent plus de temps à réaliser que d'autres et n'ont pas tous le même pointage. L'enfant doit évaluer les bénéfices et les coûts en termes de temps et de points obtenus pour gagner le plus de points possible à la fin. Des mesures du nombre total de points amassés, des items à valeur élevée essayés et des items à valeur faible essayés sont notées. La fidélité test-retest pour ce test est faible mais acceptable, avec un coefficient de 0,71. Un effet significatif de l'âge est observé sur le score obtenu à ce test chez des enfants neurotypiques.

Autocritique (insight)

Pour mesurer l'autocritique, une courbe normale hypothétique est montrée et expliquée à l'enfant par l'expérimentateur, qui lui demande de situer sa performance sur la courbe au test précédent, soit au test Fluidité catégorielle. Tel que rapporté par Kramer *et al.* (2014), une étude (Krueger *et al.*, 2011) a montré que chez un groupe d'enfants ($n = 91$) présentant ou non une atteinte cognitive, le test évaluant l'autocritique est corrélé significativement et négativement avec le score à la sous-échelle de flexibilité du *BRIEF*.

Normes sociales

Le respect des normes sociales est évalué avec un questionnaire adapté aux enfants qui est lu à l'enfant par l'expérimentateur. Le questionnaire comprend 30 items à réponse forcée

(vrai ou faux) qui font référence à des comportements inacceptables (par ex. fouiller dans ton nez en classe ou brosser tes dents une fois par semaine) et à des comportements plus acceptables (par ex. demander à ton professeur comment se rendre à la bibliothèque ou porter les mêmes souliers trois jours de suite). Deux scores sont obtenus à ce questionnaire, le score de sous-adhérence (*Break score*) et le score de sur-adhérence (*Overadhere score*). Le score de sous-adhérence indique si l'enfant a tendance à transgresser les normes sociales et le score de sur-adhérence indique si l'enfant a tendance à interpréter les normes sociales de manière trop rigide. La fidélité test-retest du questionnaire est élevée, avec un coefficient de 0,93 pour les enfants. Le score obtenu au questionnaire de normes sociales corrèle négativement et significativement avec une échelle comportementale de dysfonctions exécutives (*The Frontal Systems Behavior Scale (FrsBe)*) (Grace et Malloy, 2001) principalement au niveau de la désinhibition, de l'apathie et du dysfonctionnement exécutif global, chez 178 participants adultes, et positivement à un indice d'empathie (*Interpersonal Reactivity Index*) chez 212 participants adultes. Les concepteurs de la batterie mentionnent dans le manuel de l'examineur que le score obtenu au questionnaire de normes sociales est affecté à la fois par le fonctionnement socio-affectif et par les déficits cognitifs sur le plan des FE.

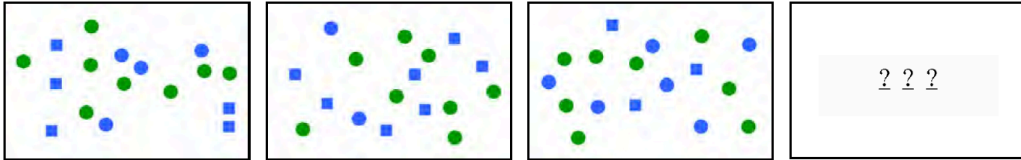
Contrôle comportemental

L'échelle comportementale évalue neuf domaines sur une échelle de Likert en quatre points (aucun, faible, modéré, sévère). Le manuel de l'examineur (Kramer *et al.*, 2014) contient une description des neuf domaines pour guider la cotation. Les neuf domaines sont l'agitation, la dépendance aux stimuli présentés, les persévérations, la diminution de l'initiation à l'action, les stéréotypies motrices, la distractibilité, le degré d'engagement social et émotionnel, l'impulsivité et les conduites sociales inappropriées. Cette échelle est remplie

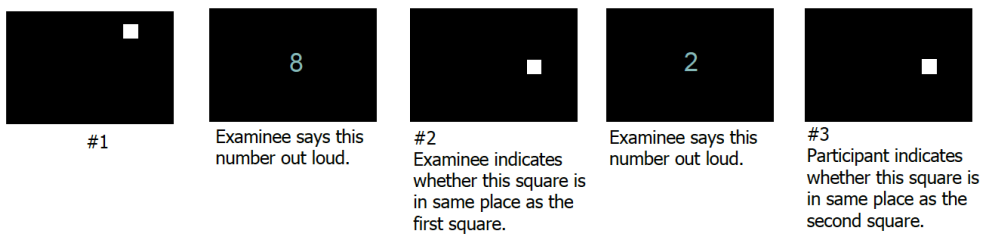
par l'expérimentateur qui note la fréquence d'apparition de comportements relatifs à chaque domaine pendant la passation de la batterie. Il n'y a cependant pas de données concernant les indices psychométriques de l'échelle comportementale dans le manuel de l'examineur.

Annexe 2 : Exemples de stimuli de la NIH-EXAMINER

Comptage de cercles



N-Back



Flanker

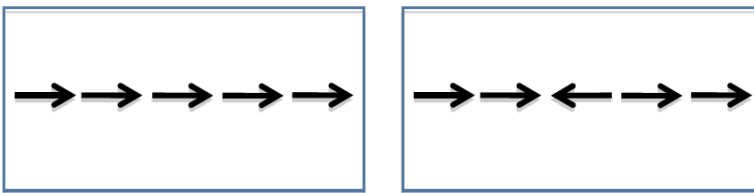


Figure 4. Example of Flanker stimuli

Dans la version pour enfants, les flèches sont remplacées par des poissons.

CPT

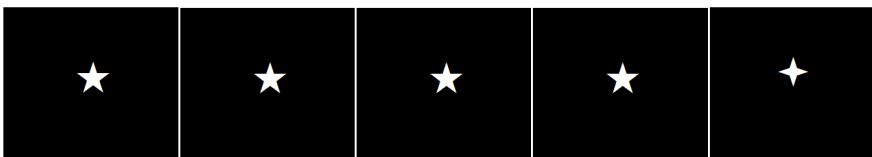


Figure 6. Example of CPT stimuli

Test d'appariement

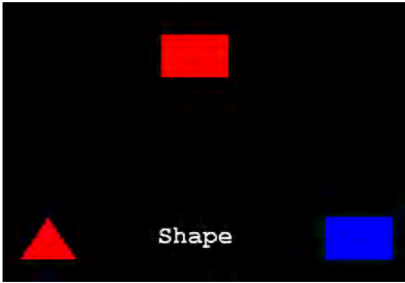
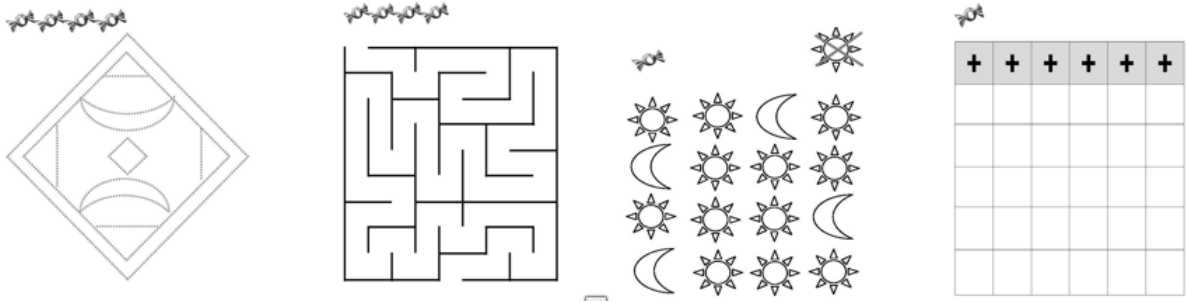


Figure 7. Example of Set Shifting stimulus

Test non-structuré



Autocritique/Insight

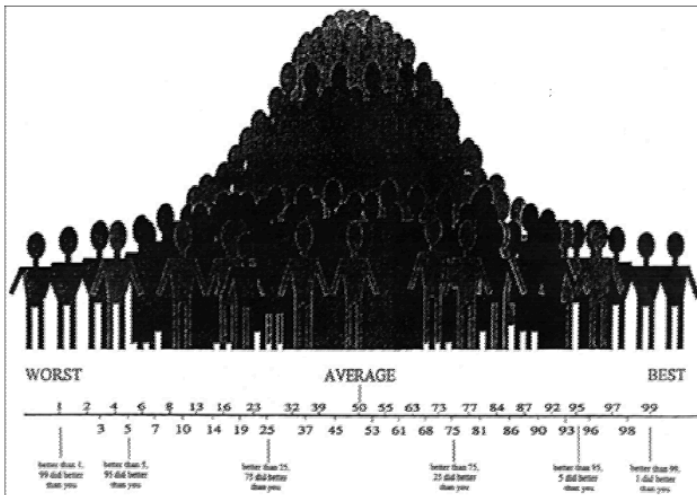


Figure 10. Normative graph for examinee self-rating

Annexe 3 : Livret de cotation de la NIH-EXAMINER

[Expérimentateur] : Je vais te donner une catégorie et je veux que tu me nommes, aussi rapidement que possible, toutes les choses qui appartiennent à cette catégorie. Par exemple, si je te dis « meubles », tu peux dire « chaise », « table », ou « bureau ». La première lettre du mot n'a pas d'importance. Maintenant je veux que tu me nommes des choses qui appartiennent à la catégorie des Animaux. Tu as une minute. Je veux que tu me dises tous les animaux auxquels tu peux penser en une minute. Tu es prêt? Vas-y!

Matériel : chronomètre 1 minute

ADMINISTRATION:

Appuyez sur la barre d'espace, le chronomètre et l'enregistrement vont se déclencher et un signal se fera entendre à la fin de la minute. Restez attentif aux réponses du participant.

Ne dites pas au participant d'inclure d'autres animaux que des mammifères. Toutefois, si le participant le demande avant le début du test ou pendant celui-ci, l'expérimentateur a la permission de dire 'oui'. Il est aussi permis de répéter les instructions ou la catégorie si le participant le demande spécifiquement.

INDICES :

1. Si le participant fait une pause de 15 secondes:
 - **Continue.**
 - **À quel autre animal peux-tu penser?**
2. Si le participant dit trois mots consécutifs qui ne commencent pas par la catégorie désignée: (Donnez cet indice une seule fois pendant la condition.)
 - **Nous utilisons présentement la catégorie « animaux ».**

Retranscription des réponses:

1.	11.	21.	31.
2.	12.	22.	32.
3.	13.	23.	33.

4.	14.	24.	34.								
5.	15.	25.	35.								
6.	16.	26.	36.								
7.	17.	27.	37.								
8.	18.	28.	38.								
9.	19.	29.	39.								
10.	20.	30.	40.								
<table> <tr> <td>Résultats :</td> <td>Corrects : _____ /</td> <td>Répétitions :</td> <td>Violation de règles :</td> </tr> <tr> <td></td> <td>40</td> <td>_____</td> <td>_____</td> </tr> </table>				Résultats :	Corrects : _____ /	Répétitions :	Violation de règles :		40	_____	_____
Résultats :	Corrects : _____ /	Répétitions :	Violation de règles :								
	40	_____	_____								

Fluidité Catégorielle: Légumes

[Expérimentateur] : Maintenant je veux que tu me nommes des choses qui appartiennent à une autre catégorie : Légumes. Tu as une minute. Je veux que tu me dises tous les légumes auxquels tu peux penser en une minute. Es-tu prêt? Vas-y !

ADMINISTRATION:

Appuyez sur la barre d'espace, le chronomètre et l'enregistrement vont se déclencher et un signal se fera entendre à la fin de la minute. Restez attentif aux réponses du participant.

Il est permis de répéter les instructions ou la catégorie si le participant le demande spécifiquement.

INDICES :

1. Si le participant fait une pause de 15 secondes:

- **Continue.**

- À quel autre légume peux-tu penser?
2. Si le participant dit trois mots consécutifs qui ne commencent pas par la lettre désignée :
(Donnez cet indice une seule fois pendant la condition.)
- Nous utilisons présentement la catégorie « légumes ».

Retranscription des réponses:

1.	11.	21.	31.
2.	12.	22.	32.
3.	13.	23.	33.
4.	14.	24.	34.
5.	15.	25.	35.
6.	16.	26.	36.
7.	17.	27.	37.
8.	18.	28.	38.
9.	19.	29.	39.
10.	20.	30.	40.

Résultats :	Corrects : _____ / 40	Répétitions : _____	Violation de règles : _____
-------------	-----------------------	---------------------	-----------------------------

Autocritique

MATÉRIEL: *Graphique de la distribution normale (Jointe)*

Après avoir complété le test de fluidité verbale :

[Expérimentateur]: J'aimerais que tu me dises comment tu penses que tu as performé à cette tâche. En comparaison avec d'autres personnes de ton âge et avec un niveau d'éducation similaire, comment penses-tu que vous vous comparez?

J'aimerais que tu me montres où tu serais sur ce graphique. Appuyer sur la barre d'espace pour présenter le graphique.

Si nous regardons un groupe de 100 personnes, nous voyons qu'il y en a très peu qui ont une très faible performance [pointez le côté gauche du graphique], et qu'il y en a très peu qui ont une performance très bonne [pointez le côté droit du graphique], et la majorité des gens se retrouvent ici, au milieu autour du 50^e percentile [pointez le centre du graphique]. Comment penses-tu avoir performé?

ADMINISTRATION:

Inscrivez l'auto-évaluation ci-dessous. Écrivez une réponse en termes de percentile, entre 1 et 100.

INDICES:

Si le participant pointe le graphique mais ne donne pas de chiffre:

- À quel nombre cela correspondrait?

Auto-Évaluation pour la fluidité verbale : _____

Test Non-Structuré

MATÉRIEL: Minuterie de la tâche informatique – 3 minutes, Page de Pratique (jointe), 3 Livrets de Stimuli (joints) et stylo.

[Expérimentateur]: Sur cette page de pratique, il y a quatre problèmes à essayer. Essaie d'abord ceux-ci.

Positionner la Page de Pratique devant le participant. Lire à voix haute les instructions sur comment compléter chaque problème et clarifier si le participant a des questions.

Voici un livret. Chaque page a quatre problèmes comme ceux que tu as fait. Comme avant, il va y avoir des problèmes où tu dois tracer, remplir des encadrés, faire des labyrinthes ou rayer la bonne figure.

Pointez le problème approprié pour chaque problème.

Pour chaque problème que tu finis, tu gagneras une, deux, trois ou quatre parties de bonbon. Regarde chaque partie de bonbon au-dessus de chaque problème.

Pointez les images de bonbon sur la première page du livret de test.

Ça, c'est le nombre de bonbons que tu vas obtenir chaque fois que tu finis un problème. Tu vas obtenir plus de bonbons pour certains problèmes et certains problèmes prennent plus de temps à finir. Tu peux faire les problèmes que tu veux et dans n'importe quel ordre. Ton but est de gagner le plus de bonbons que tu peux en 3 minutes.

Appuyer sur la barre d'espace pour montrer le chronomètre et les instructions.

ADMINISTRATION:

Placez le livret de 3 Stimuli devant le participant. Pointez les instructions si le participant semble les avoir oubliées.

Démarrez le chronomètre de l'ordinateur après avoir complété les instructions. Arrêtez le participant à 3 minutes.

Ne permettez PAS au participant de compléter un problème débuté une fois que la limite de temps est atteinte.

P1

P2

P3

P4

P5

P6

25	75	20	5	15	10	5	10	5	5	5	5	ÉC ___	FC ___
75	10	75	75	10	50	10	1	5	1	1	1	ÉE ___	FE ___
Livret 1												Total : _____	
5	75	5	25	50	1	10	10	5	5	1	5	ÉC ___	FC ___
75	20	75	75	10	15	5	1	5	1	5	1	ÉE ___	FE ___
Livret 2												Total : _____	
20	10	5	20	10	15	1	10	5	1	1	5	ÉC ___	FC ___
75	75	75	75	50	1	5	10	5	5	1	5	ÉE ___	FE ___
Livret 3												Total : _____	

Feuille réponse :

Utilisez la grille à votre gauche afin d'indiquer quels problèmes ont été essayés (E) ou complétés (C) lorsque vous corrigez les livrets. Les cases en gras indiquent les pages de chaque livret. Les items à valeur élevée (É) sont soulignés et gras, alors que les items faibles (F) sont en police normale.

Nbre d'items à haute valeur complétés : _____	Nbre d'items à faible valeur complétés : _____
Nbre d'items à haute valeur essayés : _____	Nbre d'items à faible valeur essayés : _____
Total de points obtenus : _____	

Flanker

MATÉRIEL ET MISE EN PLACE: *Voir le Manuel d'Instructions afin de mettre en place la tâche informatisée et pour plus d'informations. Utilisez les flèches Gauche et Droite du clavier (ne pas apposer d'étiquette sur les touches).*

Lisez les instructions à voix haute lorsqu'elles apparaissent à l'écran.

ESSAIS DE PRATIQUE :

[Expérimentateur] : **Regarde le poisson du milieu. Dans quelle directionnage-t-il?** Pointez le poisson au milieu. **S'il nage dans cette direction** [pointez le poisson qui nage vers la droite] **appuie sur ce bouton** [flèche de droite]. **S'il nage dans cette direction** [pointez le poisson qui nage vers la gauche] **appuie sur ce bouton** [flèche de gauche].

Essaie de répondre aussi rapidement que tu peux, sans faire d'erreurs. Essaie de garder ton attention sur la croix ("+") au centre de l'écran.

Vérifiez que les doigts du participant sont correctement placés sur les flèches gauche et droite. Encouragez le participant à garder ses doigts en place jusqu'à la fin de la tâche.

Faisons premièrement des essais de pratique. Appuyez sur la barre d'espacement pour débiter.

Effectuez les essais de pratique. Donnez de la rétroaction au participant, au besoin.

Après les essais de pratique, procédez au test. Lisez les instructions à voix haute lorsqu'elles apparaissent à l'écran.

TEST :

[Expérimentateur] :

Super! Maintenant, on peut commencer. Rappelle-toi, si le poisson au centre nage dans cette direction

[pointez le poisson qui nage vers la droite] **appuie sur ce bouton** *[flèche de droite]*. **S'il nage dans cette direction** *[pointez le poisson qui nage vers la gauche]* **appuie sur ce bouton** *[flèche de gauche]*.

Réponds aussi vite que tu peux, mais sans faire d'erreurs. Es-tu prêt?

Appuyez sur la barre d'espace pour débiter.

Test de Flexibilité cognitive

Lisez les instructions à voix haute lorsqu'elles apparaissent à l'écran.

ESSAIS DE PRATIQUE:

[Expérimentateur]: Ceci est une tâche de paires.

Regarde l'image au centre de l'écran. Si tu entends le mot FORME, tu vas former une paire entre l'image au milieu de l'écran et une des deux images dans les coins qui est de la même FORME.

Si tu entends le mot COULEUR, tu vas former une paire entre l'image au milieu de l'écran et une des deux images dans les coins qui est de la même COULEUR.

Pour trouver le même que celui-ci [pointez l'objet de référence à droite], appuie sur cette touche qui ressemble à ça [pointez le flèche de droite].

Pour trouver le même que celui-ci [pointez l'objet de référence à gauche], appuie sur cette touche qui ressemble à ça [pointez le flèche de gauche].

Premièrement, pratiquons!

*Vérifiez que les doigts du participant sont correctement placés sur les flèches gauche et droite.
Encouragez le participant à garder ses doigts en place jusqu'à la fin de la tâche.*

Essaie de répondre aussi rapidement et précisément que tu peux. Si tu fais une erreur, continue. Nous commencerons avec des essais de pratique. Appuie sur la barre d'espacement pour débiter.

Effectuez les essais de pratique. Donnez de la rétroaction au participant, au besoin.

Après les essais de pratique, procédez au test. Lisez les instructions à voix haute lorsqu'elles apparaissent à l'écran.

ESSAIS DE PRATIQUE SUPPLÉMENTAIRES:

[Expérimentateur] : Faisons d'autres essais de pratique. Les instructions sont les mêmes.

Essaie de répondre aussi rapidement et précisément que tu peux. Si tu fais une erreur, continue.

Appuie sur la barre d'espace pour débiter.

Comptage de cercles

MATÉRIEL ET MISE EN PLACE: Feuille de pointage pour enregistrer les réponses (jointe).

Ayez la feuille de pointage à votre disposition pour enregistrer les réponses. Lisez les instructions à voix haute lorsqu'elles apparaissent à l'écran.

ESSAIS DE PRATIQUE:

[Expérimentateur] : Une série d'écrans contenant des cercles bleus, des cercles verts et des carrés bleus sera présentée. Compte et souviens-toi du nombre de CERCLES BLEUS que tu verras sur chaque écran.

Compte les CERCLES BLEUS à voix haute, un à la fois, et ensuite répète le total final à voix haute immédiatement. Ceci indiquera à l'expérimentateur que tu as terminé de compter. Appuyez sur la barre d'espace pour continuer.

Un écran avec des cercles bleus et verts apparaîtra. Après que le participant ait compté chaque cercle bleu à voix haute, appuyez sur la barre d'espace pour continuer au prochain écran.

Combien de CERCLES BLEUS as-tu compté? Appuyez sur la barre d'espace pour continuer.

Maintenant, tu vas compter les CERCLES BLEUS à l'écran, et par la suite sur un autre écran. Commence à compter les cercles bleus à voix haute dès qu'ils apparaissent sur chaque écran. Souviens-toi de répéter le total final à voix haute une fois que tu as terminé de compter.

Après un certain nombre d'écrans, tu verras des points d'interrogation. Ceci indique que tu devras répéter les totaux finaux que tu as comptés sur chaque écran, dans le bon ordre.

Nous commencerons avec des essais de pratique. Appuie sur la barre d'espace pour débiter.

Effectuez les essais de pratique. Appuyez sur la barre d'espace une fois que le participant a terminé de compter les cercles à l'écran et a répété le total final. Dites au participant de commencer à compter immédiatement après que le nouvel écran apparait. Écrivez la réponse du participant sur la feuille de pointage.

Si le participant ne donne aucune bonne réponse lors des essais de pratique, arrêtez la tâche.

INDICES :

Si le participant ne répète pas le total à voix haute après avoir compté, demandez-lui seulement, lors des essais de pratique :

Souvenez-vous de répéter le total à voix haute une fois que vous avez terminé de compter à l'écran.

TEST:

[Expérimentateur] : Tu as complété les essais de pratique. Passons maintenant à la tâche. Les instructions sont les mêmes.

Compte et souviens-toi du nombre de CERCLES BLEUS sur chaque écran. Compte les CERCLES BLEUS à voix haute, un à la fois, et ensuite répète le total final à voix haute. Répète les totaux finaux que tu as comptés quand tu vois les points d'interrogation apparaître à l'écran.

Appuyez sur la barre d'espace pour débiter.

Comptage de cercles – Feuille de Pointage

ADMINISTRATION :

Inscrivez les chiffres que le participant dit à chaque écran dans la colonne Réponse. Le nombre réel de points à chaque écran est inscrit au bas de la ligne. À la fin de l'essai, inscrivez les chiffres que les participants donnent dans la colonne Rappel. Administrez tous les essais.

POINTAGE :

Donnez un point pour chaque chiffre correctement rappelé à chaque essai. Donnez un point si le nombre rappelé n'est pas le bon mais qu'il est le même que celui donné comme réponse. Inscrivez le total dans la colonne Correct. Additionnez les valeurs correctes au bas de la page (voir le Manuel pour les instructions complètes).

Pratique : a. _____ b. _____ c. _____
 5 4 7 6 2 4

Réponse	Rappel	Correct
1. _____ 3 8	1. _____	= _____
2. _____ 3 9 5	2. _____	= _____
3. _____	3. _____	= _____

5 9 3 6		
4. _____ 3 7 6 5 8	4. _____	= _____
5. _____ 3 5 6 9 4 7	5. _____	= _____
6. _____ _____	6. _____	= _____
9 3 7 8 5 6 4		
TOTAL (additionner les valeurs des colonnes ensemble pour les essais 1 à 6)		= ___/27

CPT/Go-No-Go

Matériel et Mise en place: Utilisez seulement la flèche gauche.

Lisez les instructions à voix haute au fur et à mesure qu'elles sont présentées à l'écran.

ESSAIS DE PRATIQUE:

[Expérimentateur] : Différents objets seront présentés à l'écran. Si cette étoile à cinq branches est présentée à l'écran, appuie sur la flèche de gauche. Si une autre forme est présentée, n'appuie sur aucune touche. Réponds aussi rapidement que tu peux sans faire d'erreurs. Si tu fais une erreur, continue. Nous commencerons avec des essais de pratique.

Vérifiez que le doigt du participant soit placé de manière appropriée sur la flèche gauche. Encouragez le participant à garder son doigt en place jusqu'à ce que la tâche soit terminée.

Appuyez sur la barre d'espace pour continuer.

Effectuez les essais de pratique en donnant des rétroactions si nécessaire. Les instructions sont les mêmes pour chaque essai. Une fois que la pratique est terminée, procédez à la tâche.

TEST:

[Expérimentateur] : Procédons maintenant au test en tant que tel. Appuie sur la barre d'espace lorsque tu es prêt à débiter.

N-BACK

Matériel et Mise en place: Utilisez les flèches gauche et droite.

Lisez les instructions à voix haute au fur et à mesure qu'elles sont présentées à l'écran.

[Expérimentateur] : Tu verras une série de carrés blancs apparaître à l'écran. Souviens-toi de l'emplacement de chaque carré lorsqu'il apparaît. Tu vas le comparer à l'emplacement du prochain carré que tu verras.

Si l'emplacement d'un carré est le même que celui du carré précédent, appuie sur la touche de GAUCHE pour dire OUI. Si l'emplacement d'un carré n'est pas le même que celui du carré précédent, appuie sur la touche de DROITE pour dire NON. Entre chaque carré, un nombre va apparaître au milieu de l'écran. Dis ce nombre à voix haute. Commençons par des essais de pratique. Appuie sur la barre d'espace pour débiter.

Souviens-toi de l'emplacement de ce carré, afin de pouvoir le comparer à l'emplacement du prochain carré que tu verras. Appuyez sur la barre d'espace pour débiter. Le chiffre apparaît au milieu de l'écran. Dis ce nombre à voix haute. L'écran change et le prochain carré apparaît à l'écran. Est-ce le même emplacement que celui juste avant? Si OUI, appuie sur la TOUCHE GAUCHE. Si NON, appuie sur la TOUCHE DROITE.

Demandez au participant de répondre.

Lisez les instructions pour les trois premiers carrés et continuez à lire si le participant semble avoir de la difficulté à comprendre la tâche. Donnez de la rétroaction si nécessaire. Si le participant performe bien, les essais de pratique vont commencer.

ESSAIS DE PRATIQUE:

[Expérimentateur] : Essayons quelques carrés de pratique supplémentaires. Tu ne recevras pas d'indications ou de rétroaction pour cette ronde. Compare l'emplacement de chaque carré à celui juste avant. Dis le nombre à voix haute quand tu le vois. Appuie sur la barre d'espace pour débiter.

Verbalisez les instructions et donnez des indications si nécessaire pour aider le participant à comprendre la tâche.

ESSAIS DE PRATIQUE SUPPLÉMENTAIRES:

[Expérimentateur] : Essayons quelques carrés de pratique supplémentaires. Tu ne recevras pas d'indications ou de rétroaction pour cette ronde. Chaque carré sera présenté moins longtemps. Réponds aussi rapidement que tu peux sans faire d'erreurs. Appuie sur la barre d'espace pour débiter.

Effectuez les essais de pratique. Donnez de la rétroaction si nécessaire.

Une fois que la pratique est terminée, passez à la tâche.

TEST:

[Expérimentateur] : Il est maintenant temps de débiter la tâche. Réponds aussi rapidement que tu peux sans faire d'erreurs. Appuie sur la barre d'espace lorsque tu es prêt à débiter.

Évaluation des normes sociales

Administrez à tous les participants. Pour les participants de 14 ans et plus, administrez l'Évaluation des Normes Sociales pour Adultes. Pour les participants de moins de 14 ans, administrez l'Évaluation des Normes Sociales pour Enfants.

Instructions pour l'Évaluation des Normes Sociales pour Adultes :

[Expérimentateur] : Voici une liste de comportements que peut avoir une personne. Veuillez décider s'il est socialement acceptable et approprié de faire ces choses dans la culture dominante du Canada et répondez par oui ou par non. Voyez ces questions comme vous les appliqueriez à un étranger ou à une connaissance, pas à un bon ami ou à un membre de la famille.

Instructions pour l'Évaluation des Normes Sociales pour Enfants:

[Expérimentateur] : Voici une liste de comportements que peut avoir un autre enfant. Décide si cela dérangerait les autres si tu faisais chacune de ces choses. Fais comme si tous les enfants de ta classe et ton professeur te voyaient faire ces choses, et décide s'ils penseraient que c'est correct.

EST-CE QUE LES AUTRES PENSERAIENT QUE C'EST CORRECT DE :		
1. FOUILLER DANS TON NEZ EN CLASSE	OUI	NON
2. DEMANDER À TON PROFESSEUR COMMENT SE RENDRE À LA BIBLIOTHÈQUE	OUI	NON
3. DIRE À UN AUTRE ENFANT QUE TU NE L'AIMES PAS	OUI	NON
4. DEMANDER À TON PROFESSEUR DE T'EMMENER AU MAGASIN	OUI	NON
5. TE MOUCHER À BORD DE L'AUTOBUS	OUI	NON
6. EMPRUNTER LE NOUVEAU CHANDAIL DE TON AMI SANS PERMISSION	OUI	NON
7. BROSSER TES DENTS UNE FOIS PAR SEMAINE	OUI	NON
8. AMENER UN LIVRE DE LA BIBLIOTHÈQUE À LA MAISON	OUI	NON
9. PORTER LES MÊMES SOULIERS TROIS JOURS DE SUITE	OUI	NON

10. GARDER DE L'ARGENT QUE TU TROUVES SUR LE TROTTOIR	OUI	NON
11. DIRE À UN AUTRE ENFANT QUE SES CHEVEUX SONT TRANGES	OUI	NON
12. RIRE DE TOI-MÊME QUAND TU TRÉBUCHES ET QUE TU TOMBES	OUI	NON
13. BOTTER UN BALLON TRÈS FORT	OUI	NON
14. RIRE À VOIX HAUTE PENDANT UN FILM AU CINÉMA	OUI	NON
15. POUSSER UN AUTRE ENFANT POUR QUE TU ARRIVES AUX BALANÇOIRES EN PREMIER	OUI	NON
16. JETER LES DÉCHETS DE TON DÎNER SUR LE PLANCHER	OUI	NON
17. DEMANDER À UN AUTRE ENFANT S'IL VEUT VENIR CHEZ TOI	OUI	NON
18. MANGER LE DERNIER BISCUIT DE LA BOÎTE	OUI	NON
19. AMENER LES CRAYONS DE LA CLASSE À LA MAISON	OUI	NON
20. TE DÉPÊCHER POUR T'ASSEOIR DANS LA RANGÉE D'EN AVANT LA PREMIÈRE JOURNÉE D'ÉCOLE	OUI	NON
21. PRENDRE UN BAIN OU UNE DOUCHE À CHAQUE JOUR	OUI	NON
22. RIRE D'UN AUTRE ENFANT QUAND IL TRÉBUCHE ET TOMBE	OUI	NON
23. PARLER FORT PENDANT UN FILM AU CINÉMA	OUI	NON
24. PARLER AVEC LA BOUCHE PLEINE	OUI	NON
25. GARDER LE JEU QU'UN CAMARADE A LAISSÉ DANS LA COUR D'ÉCOLE	OUI	NON

26. DIRE À UN AUTRE ENFANT QU'IL EST INTELLIGENT	OUI	NON
27. BOTTER LES POUBELLES DE L'ÉCOLE	OUI	NON
28. JETER LES RESTANTS DE TON DÎNER	OUI	NON
29. UTILISER UNE CUEILLÈRE POUR MANGER DU MACARONI AU FORMAGE	OUI	NON
30. PORTER LE MÊME CHANDAIL TROIS JOURS DE SUITE	OUI	NON

Collecte de Données / Notes sur la Qualité de la Recherche

Problèmes de Qualité :			
<input type="checkbox"/> Aucun, les données sont valides (passez aux notes)	<input type="checkbox"/> Difficultés motrices	<input type="checkbox"/> Français langue seconde	<input type="checkbox"/> Problèmes de Comportement
	<input type="checkbox"/> Difficultés verbales	<input type="checkbox"/> Éducation minimale	
<input type="checkbox"/> Oui (choisissez jusqu'à 3 problèmes)	<input type="checkbox"/> Difficultés d'audition	<input type="checkbox"/> Manque d'effort	<input type="checkbox"/> Autres (décrivez dans les notes)
	<input type="checkbox"/> Difficultés visuelles	<input type="checkbox"/> Informateur non fiable	
Notes sur la Qualité des Données :			

Échelle Comportementale

INSTRUCTIONS:

À compléter par l'Expérimentateur. Suivant l'administration de la batterie au participant, indiquez la présence des types de comportements suivants. « Aucun » indique l'absence complète de ce type de comportement. Lorsque le comportement est présent, notez « léger », « modéré » ou « sévère » selon le degré auquel la passation des tests ou les échanges interpersonnels sont perturbés, ou la mesure selon laquelle les comportements dévient des normes sociales acceptées.

Veillez encercler la réponse reflétant le mieux le comportement du participant.

DESCRIPTION DES ITEMS ET LISTE DE COMPORTEMENTS:

1. Agitation	Aucun	Faible	Modéré	Sévère
2. Dépendance aux stimuli	Aucun	Faible	Modéré	Sévère
3. Persévération	Aucun	Faible	Modéré	Sévère
4. Initiative diminuée	Aucun	Faible	Modéré	Sévère
5. Stéréotypies motrices	Aucun	Faible	Modéré	Sévère
6. Distractibilité	Aucun	Faible	Modéré	Sévère
7. Manque d'engagement social / émotionnel	Aucun	Faible	Modéré	Sévère
8. Impulsivité	Aucun	Faible	Modéré	Sévère
9. Conduites sociales inappropriées	Aucun	Faible	Modéré	Sévère

Descriptions:

Directives Générales

Les expérimentateurs devraient restreindre leur évaluation aux comportements qu'ils ont observés, et ne pas utiliser cette échelle afin de refléter des comportements décrits par des soignants, informateurs ou autres professionnels de la santé n'ayant pas été observés directement par l'expérimentateur. Les expérimentateurs devraient inclure tous les comportements, sans égard au contexte. Ainsi, même si les comportements observés pendant l'expérimentation représentent l'essentiel des données, les expérimentateurs peuvent aussi noter des comportements observés dans d'autres situations, comme la salle d'attente et le trajet entre la salle d'attente et la salle d'expérimentation.

Dans certains cas, les expérimentateurs devront décider à quelle catégorie un comportement en particulier s'applique. Par exemple, un participant ramassant de manière répétée un crayon et gribouillant sur les feuilles de test pourrait être vu comme un comportement de persévération, de dépendance aux stimuli ou de stéréotypie motrice. Il est important que les expérimentateurs sélectionnent **seulement une** catégorie pour chaque comportement observé. Typiquement, les expérimentateurs vont utiliser leur meilleur jugement clinique afin de déterminer la catégorie la plus appropriée. Certaines suggestions afin de déterminer la bonne catégorie sont mentionnées ci-dessous. Utilisez l'espace dans la moitié inférieure de la page de notation afin de décrire le comportement. Ceci est particulièrement important pour les sujets qui seront évalués de manière longitudinale, dans le cas où un expérimentateur différent verrait un participant lors d'une visite ultérieure.

DÉTERMINER LA SÉVÉRITÉ:

Comme règle générale, la sévérité du comportement observé devrait indiquer le degré auquel la passation des tests ou les échanges interpersonnels sont perturbés, ou la mesure selon laquelle les comportements dévient des normes sociales acceptées. **Faible** réfère à une occurrence peu fréquente du comportement ou si le comportement observé est présent, mais relativement négligeable. Les sujets sont facilement redirigeables et il n'y a pas ou peu d'impact sur la qualité de la passation. Une notation **Modéré** indique que l'occurrence du comportement commence à contrevenir à la qualité des données et à la performance aux tests neuropsychologiques. Les sujets sont moins facilement redirigés. Par exemple, si un sujet ne considère pas les indices sociaux de l'expérimentateur et continue de discuter de sujets qui sont inappropriés dans un contexte clinique, l'élément « Conduites Sociales Inappropriées » serait modérée. Une notation Sévère indique que le comportement se produit très fréquemment lors de la situation d'expérimentation. Par exemple, la distractibilité serait sévère quand il est très difficile de rediriger le participant vers la tâche, au point de questionner la validité de la tâche. Des exemples additionnels sont fournis plus bas.

COMPORTEMENTS SPÉCIFIQUES :

- 1. Agitation:** L'agitation peut impliquer des comportements inappropriés verbaux (crier, jurer) ou physiques (mouvements corporels répétitifs, frapper ou lancer des objets) et peuvent être agressifs ou non dans leur nature (gigoter la jambe versus donner des coups de pied). De l'agitation légère peut se manifester en étant anxieux face à l'évaluation, en argumentant ou en se plaignant de l'expérimentation. Des niveaux modérés peuvent se manifester par des comportements perturbant l'évaluation sans être nuisibles et ne peuvent pas facilement être redirigés. De l'agitation sévère inclut des comportements physiques qui mettent le participant ou les autres en danger (par ex. frapper, pousser, graffigner, lancer des objets, mordre ou donner des coups de pieds) ou de l'agression verbale extrême (par ex. hurler ou jurer fort).
- 2. Dépendance aux stimuli :** C'est une réponse inappropriée à un stimulus saillant de l'environnement. De tels comportements peuvent inclure la lecture non-sollicitée d'un

texte aux alentours, de la dépendance à l'environnement (par ex. ramasser un crayon de sur la table et crier, manger la nourriture dans l'assiette de quelqu'un d'autre), une attention excessive aux objets qui ne sont pas pertinents (par ex. ramasser des objets sur le plancher), des écholalies ou, pendant le test, écrire ou dessiner sur le matériel ou sur un stimulus adjacent. Une seule occurrence de ce type de comportement peut justifier l'inscription d'un niveau Léger. Des exemples d'un niveau modéré seraient par exemple une fixation complète de l'attention sur un stimulus ou le fait d'être distrait par des stimuli adjacents et que ces comportements aient clairement pour conséquence une détérioration de la performance. Les comportements dépendants de l'environnement surviennent souvent et le participant ne peut pas être redirigé. Le participant dont la dépendance à l'environnement rend la complétion de la tâche très difficile recevrait la mention sévère.

3. **Persévération :** La persévération est une persistance inappropriée et non-intentionnelle d'un comportement et peut-être observée de manière comportementale, lors du *testing*, dans le discours et dans la conversation. Tel qu'observé dans le *testing*, le comportement persistant peut être (1) une répétition de réponses générées précédemment dans une tâche, ou (2) une répétition d'une réponse appropriée dans une tâche ou condition précédente. Tel qu'observé dans le discours et la conversation, le participant peut sembler être bloqué sur une idée ou retourner de manière persistante à une idée ou une histoire énoncée précédemment. (Quelques réponses répétées pendant le *testing* ne sont pas rares chez les individus normaux, donc ce comportement devrait seulement être noté lorsqu'il est présent à une fréquence qui dévie des attentes normales.) Des répétitions qui

excèdent légèrement les attentes normales seraient évaluées comme légères. Une notation modérée serait donnée dès le moment où le participant persévère dans sa réponse sans réponses intermédiaires (par ex. nommer exactement le même mot deux fois ou plus de suite dans la catégorie de fluidité verbale) ou retourne vers des réponses données précédemment (donner au deuxième bloc de la fluidité verbale des réponses qui commencent par la lettre L tel que demandé dans le premier bloc), ou qui fait plusieurs persévérations de plusieurs types. Une tendance à la persévération envahissante menant à l'interruption de la passation des tests serait considérée comme sévère.

4. **Initiative diminuée :** Des comportements indiquant une initiative diminuée peuvent inclure un délai entre l'initiation motrice ou verbale et la fin des instructions, le besoin de se faire inciter avant d'initier une réponse, et une faible motivation à bien performer. Ceci pourrait aussi inclure des instances où un participant arrête au milieu de la tâche et demande de se faire encourager à continuer. Note : Si l'initiative diminuée se produit seulement dans le contexte d'un manque d'engagement social ou de distractibilité, notez-le dans ces domaines plutôt que dans l'initiative diminuée.

5. **Stéréotypies motrices :** Ce sont des comportements persistants et répétitifs sans but précis. Cela peut inclure tapoter des doigts ou gigoter la jambe sans but apparent, fouiller dans les poches ou les tiroirs, se toucher la peau ou rajuster ses vêtements, manipuler des boutons, enlever et mettre des bijoux ou des vêtements de manière répétée, faire des bruits avec sa bouche ou répéter des phrases qui n'ont pas de valeur communicative. Note : Si un comportement est codé comme étant une stéréotypie

motrice, le même comportement ne devrait pas être codé comme une persévération.

- 6. Distractibilité :** L'attention est facilement dérangée par des stimuli externes (gens, objets, etc.), le bruit externe ou les pensées internes. Ceci est observable lorsque le participant perd le fil de sa pensée, a besoin de plusieurs rappels des instructions, a besoin que l'on redirige son attention sur la tâche, ou fait preuve de discours ou de pensées tangentiels. Ces comportements peuvent être considérés comme faibles si la distractibilité est occasionnelle ou qu'elle n'interfère pas significativement avec une collecte de données valides, modérés si le participant a besoin de plusieurs rappels et redirections, mais est toujours capable de finir l'évaluation, et sévères si la distractibilité interfère avec la collecte de données au point d'en questionner la validité.

- 7. Manque d'Engagement Social / Émotionnel :** Ceci réfère à l'impression qu'a l'expérimentateur de signes de manque d'intérêt social, d'interrelation ou de chaleur personnelle. Des exemples de ce type de comportements peuvent inclure un manque de contacts visuels, un excès de contact visuel (par ex. fixer l'expérimentateur ou un autre individu), un manque de sourire ou une gamme réduite d'expressions faciales (n'étant pas reliée à une maladie de Parkinson), un manque de conscience de comment est-ce que le comportement d'un individu peut en affecter un autre ou un manque d'intérêt pour l'autre. D'autres comportements peuvent inclure un manque de spontanéité et le fait de ne pas engager la conversation. La distinction principale entre le manque d'engagement social et les conduites sociales inappropriées est que le manque d'engagement social reflète typiquement une absence de comportement (par ex.

empathie, chaleur) alors que les conduites sociales inappropriées reflètent la présence distincte d'un comportement inapproprié (par ex. remarques ou toucher inapproprié).

8. Impulsivité : Ceci réfère au fait d'agir avec une réflexion ou une patience insuffisante. Commencer une tâche avant que l'expérimentateur ait terminé les instructions, répondre à seulement une partie des instructions (par ex. la première partie d'une série d'instructions), une approche peu soignée face au travail, finir trop rapidement et interrompre l'expérimentateur sont des comportements potentiels de la catégorie Impulsivité. Une notation faible référerait au fait que le participant travaille trop rapidement, sans vérifier son travail, semble s'ennuyer lorsqu'il résout des problèmes complexes, ou débute un item d'un test avant que l'expérimentateur n'ait terminé les instructions ou sans planification adéquate. Pour équivaloir à une notation modérée, le participant peut répondre sans considération de l'exactitude, débiter les items du test avant que l'expérimentateur n'ait fini les instructions ou sans planification adéquate à plus d'une occasion, ou a tendance à faire ou à dire les choses rapidement, sans planification. Un exemple d'impulsivité sévère serait quand la tendance du participant à agir sans réfléchir interrompt la collecte de données ou les échanges interpersonnels, ou dévie de la norme de manière notable.

9. Conduites sociales inappropriées : Réfère à des comportements qui ne sont pas adéquats dans un contexte professionnel. Ceci peut se traduire, par exemple, par parler à des étrangers dans l'ascenseur, poser des questions personnelles ou visant les compétences de l'expérimentateur, ne pas considérer ou violer l'espace personnel (par

ex. toucher l'expérimentateur de manière inappropriée ou excessive), ou par le manque de réponse aux indices sociaux (par ex. lorsque les autres essaient de terminer une conversation). Les sujets inappropriés, tabous ou les commentaires à connotation sexuelle, des blagues ou des opinions qui peuvent être offensantes pour les autres, ou une pauvre hygiène (par ex. mauvaises odeurs, vêtements tachés, déchirés ou inappropriés) peuvent aussi être inclus dans cette catégorie. Des comportements physiques tels que des flatulences, toucher à des parties du corps privées, roter et cracher peuvent aussi être considérés.

Références

- Akshoomoff, N., Brown, T. T., Bakeman, R., & Hagler Jr, D. J. (2018). Developmental differentiation of executive functions on the NIH Toolbox Cognition Battery. *32*(7), 777.
- Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *8*(2), 71-82.
- Anderson, V., Jacobs, R., & Anderson, P. J. (2010). *Executive functions and the frontal lobes: A lifespan perspective* : Psychology Press.
- Anthoine, E., Moret, L., Regnault, A., Sébille, V., & Hardouin, J.-B. (2014). Sample size used to validate a scale: a review of publications on newly-developed patient reported outcomes measures. *12*(1), 2.
- Arafat, S., Chowdhury, H. R., Qusar, M., & Hafez, M. (2016). Cross cultural adaptation & psychometric validation of research instruments: A methodological review. *5*(3), 129-136.
- Ardila, A. (2005). Cultural values underlying psychometric cognitive testing. *15*(4), 185.
- Baddeley, A. (1986). Oxford psychology series, No. 11. Working memory. New York, NY, US. In : Clarendon Press/Oxford University Press.
- Baddeley, A. (1996). Exploring the central executive. *49*(1), 5-28.
- Baddeley, A. (2000). The episodic buffer: a new component of working memory? , *4*(11), 417-423.
- Baddeley, A. (2012). Working memory: theories, models, and controversies. *63*, 1-29.
- Baddeley, A., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47-89): Elsevier.
- Bandura, A. (2014). Social cognitive theory of moral thought and action. In *Handbook of moral behavior and development* (pp. 69-128): Psychology Press.
- Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *121*(1), 65.

- Barkley, R. A. (2012). *Executive functions: What they are, how they work, and why they evolved* : Guilford Press.
- Baron, I. S. (2007). Behavioural Assessment of the Dysexecutive Syndrome for Children (BADS-C) by Emslie, H., Wilson, FC, Burden, V., Nimmo-Smith, I., & Wilson, BA (2003) London, UK: Harcourt Assessment/The Psychological Corporation. BADS (-C) May Be Good.
- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *25*(24), 3186-3191.
- Beauchamp, M., & Dooley, J. (2013). *Exploring the cognitive and affective predictors of moral reasoning in adolescence using the So-Moral task*. Paper presented at the biennial meeting of the Society for Research in Child Development, Seattle.
- Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *21*(4), 327-336.
- Blair, C., & Diamond, A. (2008). Biological processes in prevention and intervention: The promotion of self-regulation as a means of preventing school failure. *20*(3), 899-911.
- Burgess, P. W., & Shallice, T. (1997). The relationship between prospective and retrospective memory: Neuropsychological evidence. *5*, 249-256.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *56*(2), 81.
- Carlson, S. M., & Wang, T. S. (2007). Inhibitory control and emotion regulation in preschool children. *22*(4), 489-510.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. 2nd. In : Hillsdale, NJ: Erlbaum.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *52*(4), 281.
- Delis, D., Kramer, J., Kaplan, E., & Ober, B. (2000). *California Verbal Learning Test* . San Antonio, TX: The Psychological Corporation. In : Harcourt Assessment Company.
- Delis, D. C., Jacobson, M., Bondi, M. W., Hamilton, J. M., & Salmon, D. P. (2003). The myth of testing construct validity using factor analysis or correlations with normal or mixed clinical populations: Lessons from memory assessment. *9*(6), 936-946.

- Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive function system: examiners manual* : Psychological Corporation.
- Dumont, L. (2017). *Évaluation systématique des effets de la tDCS sur le DLPFC et applications en technologies de l'information*. (Thèse de Doctorat), Université de Montréal, Retrieved from <https://papyrus.bib.umontreal.ca/xmlui/handle/1866/20616>
- Dupaul, G., Power, T., Anastopoulos, A., & Reid, R. (1998). ADHD rating scale-IV, Guilford.
- Emslie, H., Burden, V., Nimmo-Smith, I., Wilson, B. A., & Wilson, C. (2003). *Behavioural assessment of the dysexecutive syndrome for children* : Thames Valley Test Company.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *41*(4), 1149-1160.
- Gershon, R. C., Cella, D., Fox, N. A., Havlik, R. J., Hendrie, H. C., & Wagster, M. V. (2010). Assessment of neurological and behavioural function: the NIH Toolbox. *9*(2), 138-139.
- Gioia, G. A., & Isquith, P. K. (2004). Ecological assessment of executive function in traumatic brain injury. *25*(1-2), 135-158.
- Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). Test review behavior rating inventory of executive function. *6*(3), 235-238.
- Goldstein, S., & Naglieri, J. A. (2013). *Handbook of executive functioning* : Springer Science & Business Media.
- Grace, J., & Malloy, P. F. (2001). *Frontal systems behavior scale: FrSBe* : Psychological Assessment Resources Lutz, FL.
- Heaton, R. K. (1981). *A manual for the Wisconsin card sorting test* : Western Psychological Services.
- Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). *Wisconsin Card Sorting Test (WCST): Manual: Revised and Expanded* : Psychological Assessment Resources (PAR).
- Hogan, T. P. (2017). *Introduction à la psychométrie* : Chenelière éducation.
- Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: a review of our current understanding. *17*(3), 213-233.
- Kaplan, E., Goodglass, H., & Weintraub, S. (1983). The Boston naming test. 2nd.

- Kennepohl, S. (1999). Toward a cultural neuropsychology: An alternative view and a preliminary model. *41*(3), 365-380.
- Korkman, M., Kirk, U., & Kemp, S. (2014). *NEPSY-II* : Pearson.
- Kramer, J. H., Mungas, D., Possin, K. L., Rankin, K. P., Boxer, A. L., Rosen, H. J., . . . Widmeyer, M. (2014). NIH EXAMINER: conceptualization and development of an executive function battery. *20*(1), 11-19.
- Krueger, C. E., Rosen, H. J., Taylor, H. G., Espy, K. A., Schatz, J., Rey-Casserly, C., & Kramer, J. H. (2011). Know thyself: real-world behavioral correlates of self-appraisal accuracy. *25*(5), 741-756.
- Larrabee, G. (2003). Lessons on measuring construct validity: A commentary on Delis, Jacobson, Bondi, Hamilton, and Salmon. *9*(6), 947-953.
- Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *21*(1), 59-80.
- Lewis, C., & Carpendale, J. (2009). Introduction: Links between social interaction and executive function. *2009*(123), 1-15.
- Lezak. (1995). *Neuropsychological assessment*, Oxford Univ.
- Lezak, M. D., Howieson, D. B., Loring, D. W., & Fischer, J. S. (2004). *Neuropsychological assessment* : Oxford University Press, USA.
- Luria, A. (1973). *The working brain: An introduction to neuropsychology*. .
- Luria, A. R. (1973). The frontal lobes and the regulation of behavior. In *Psychophysiology of the frontal lobes* (pp. 3-26): Elsevier.
- Luria, A. R. (1980). Disturbances of higher cortical functions with lesions of the frontal region. In *Higher cortical functions in man* (pp. 246-365): Springer.
- Lussier, F., & Flessas, J. (2009). *Neuropsychologie de l'enfant: troubles développementaux et de l'apprentissage* (Dunod Ed.).
- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *4*(1), 84.
- Manly, T., Robertson, I. H., Anderson, V., Nimmo-Smith, I., Lussier, F., & Flessas, J. (2006). *TEA-Ch: Test d'Evaluation de l'Attention Chez l'enfant* : Éditions du Centre de psychologie appliquée (ECPA).

- Markova, I., & Berrios, G. (1992). The assessment of insight in clinical psychiatry: a new scale. *86*(2), 159-164.
- Mazeau, M., & Pouhet, A. (2014). *Neuropsychologie et troubles des apprentissages chez l'enfant: du développement typique aux dys* : Elsevier Masson.
- Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *21*(1), 8-14.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *41*(1), 49-100.
- Morand-Beaulieu, S., Leclerc, J., Valois, P., Lavoie, M., O'Connor, K., & Gauthier, B. (2017). A review of the neuropsychological dimensions of Tourette syndrome. *7*(8), 106.
- Moriguchi, Y. J. F. i. p. (2014). The early development of executive function and its relation to social interaction: a brief review. *5*, 388.
- Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *81*(2), 111-121.
- Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: views from cognitive and personality psychology and a working inhibition taxonomy. *126*(2), 220.
- Nigg, J. T., Willcutt, E. G., Doyle, A. E., & Sonuga-Barke, E. (2005). Causal heterogeneity in attention-deficit/hyperactivity disorder: do we need neuropsychologically impaired subtypes? , *57*(11), 1224-1230.
- Norman, D. A., & Shallice, T. (1986). Attention to action. In *Consciousness and self-regulation* (pp. 1-18): Springer.
- Packwood, S., Hodgetts, H. M., & Tremblay, S. (2011). A multiperspective approach to the conceptualization of executive functions. *33*(4), 456-470.
- Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *37*(1), 51-87.
- Peter, G. (2010). Développement des processus d'inhibition et de flexibilité du système superviseur attentionnel.
- Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *13*(1), 25-42.

- Possin, K. L., LaMarre, A. K., Wood, K. A., Mungas, D. M., & Kramer, J. H. (2014). Ecological validity and neuroanatomical correlates of the NIH EXAMINER executive composite score. *20*(1), 20-28.
- Regard, M., Strauss, E., & Knapp, P. (1982). Children's production on verbal and non-verbal fluency tasks. *55*(3), 839-844.
- Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1994). The test of everyday attention (TEA).
- Rueda, M. R., & Paz-Alonzo, P. (2013). Executive function and emotional development. 17.
- Schatz, J., Stancil, M., Katz, T., & Sanchez, C. E. (2014). EXAMINER executive function battery and neurologic morbidity in pediatric sickle cell disease. *20*(1), 29-40.
- Sergeant, J. A., Geurts, H., & Oosterlaan, J. (2002). How specific is a deficit of executive functioning for attention-deficit/hyperactivity disorder? , *130*(1-2), 3-28.
- Shallice. (1982). Specific impairments of planning. *298*(1089), 199-209.
- Shallice, & Burgess. (1996). The domain of supervisory processes and temporal organization of behaviour. *351*(1346), 1405-1412.
- Shallice, T., Marzocchi, G. M., Coser, S., Del Savio, M., Meuter, R. F., & Rumati, R. I. (2002). Executive function profile of children with attention deficit hyperactivity disorder. *21*(1), 43-71.
- Simonds, J., Kieras, J. E., Rueda, M. R., & Rothbart, M. K. (2007). Effortful control, executive attention, and emotional regulation in 7–10-year-old children. *22*(4), 474-488.
- Sonuga-Barke, E. (2003). The dual pathway model of AD/HD: an elaboration of neuro-developmental characteristics. *27*(7), 593-604.
- Sonuga-Barke, E. (2005). Causal models of attention-deficit/hyperactivity disorder: from common simple deficits to multiple developmental pathways. *57*(11), 1231-1238.
- Strauss, E., Sherman, E. M., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* : American Chemical Society.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *18*(6), 643.
- Stuss, D. T., & Benson, D. F. (1986). *The frontal lobes* : Raven Pr.
- Stuss, D. T., Shallice, T., Alexander, M. P., & Picton, T. W. (1995). A multidisciplinary approach to anterior attentional functions. *769*(1), 191-212.

- Suchy, Y. (2009). Executive functioning: Overview, assessment, and research issues for non-neuropsychologists. *37*(2), 106-116.
- Suchy, Y. (2015). *Executive functioning: A comprehensive guide for clinical practice* : Oxford University Press.
- Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner review: Do performance-based measures and ratings of executive function assess the same construct? , *54*(2), 131-143.
- Tranel, D. (1994). Development of the concept of 'executive function' and its relationship to the frontal lobes.
- Vallerand, R. J. (1989). Vers une méthodologie de validation trans-culturelle de questionnaires psychologiques: Implications pour la recherche en langue française. *30*(4), 662.
- Wasserman, T., & Wasserman, L. D. (2013). Toward an integrated model of executive functioning in children. *2*(2), 88-96.
- Wechsler, D. (2004). *The Wechsler intelligence scale for children*. London: Pearson Assessment.
- Wechsler, D. (2008). *Wechsler Adult Intelligence Scale—Fourth Edition (WAIS–IV)* : San Antonio, TX: The Psychological Corporation.
- Wechsler, D., & Naglieri, J. (2006). Wechsler nonverbal scale of ability. PsychCorp Edition, A Brand of Harcourt Assessment. In : Pearson, San Antonio, TX.
- Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A., Chevalier, N., & Espy, K. A. (2011). The structure of executive function in 3-year-olds. *108*(3), 436-452.
- Yong, A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *9*(2), 79-94.
- Zelazo, P. D., Carter, A., Reznick, J. S., & Frye, D. (1997). Early development of executive function: A problem-solving framework. *1*(2), 198.