

Université de Montréal

**Prévalidation d'un test de fluidité graphique sous forme d'application mobile pour les enfants  
de 6 à 12 ans**

Par  
Frédérique Bolduc

Département de psychologie, Faculté des arts et des sciences

Essai présenté en vue de l'obtention du grade de doctorat en psychologie, option neuropsychologie  
clinique (D.Psy.)

Août 2023

© Frédérique Bolduc, 2023

## Table des matières

Résumé.....	ii
Abstract.....	iii
Liste des tableaux.....	iv
Liste des figures.....	v
Liste des abréviations.....	vi
Introduction.....	1
Méthodologie.....	6
Participants.....	6
Procédure.....	6
Matériel.....	7
Analyse des données et prédictions.....	11
Résultats.....	13
Description de l'échantillon.....	13
Performance, facteurs sociodémographiques et QI.....	14
Fidélité test-retest.....	17
Équivalence.....	18
Validité de construit.....	19
Performances au AppFLO.....	21
Fidélité et validité.....	24
Limites.....	29
Conclusion.....	30
Références.....	32

## Résumé

La technologie évolue à grande vitesse, mais la profession de neuropsychologue ne semble pas s'être adaptée à ces changements. Les outils informatisés sont encore peu intégrés à la pratique. Pourtant, une utilisation plus fréquente d'outils numériques aurait plusieurs avantages, tels qu'une plus grande fidélité et validité des tests, une économie de temps et d'argent, ainsi qu'une plus grande accessibilité des services neuropsychologiques, entre autres. Dans cette optique, nous avons élaboré un test de fluidité graphique sous forme d'application mobile inspiré du Five-Point Test, le test AppFLO. L'objectif de ce projet était d'effectuer une prévalidation du test AppFLO auprès d'enfants de 6 à 12 ans. Une batterie de tests et de questionnaires a été administrée à 30 enfants neurotypiques dans le but de vérifier la fidélité et la validité de l'outil. Les résultats indiquent que la performance au AppFLO s'améliore avec l'âge, le niveau scolaire et le fonctionnement intellectuel et est indépendante du sexe et du niveau d'éducation des parents. De plus, les résultats indiquent que l'outil présente une bonne fidélité test-retest. Des corrélations de modérées à fortes entre le AppFLO, le Five-Point Test et d'autres tests mesurant les fonctions exécutives laissent entrevoir une bonne validité de construit. En conclusion, cette prévalidation suggère que le test AppFLO est un outil pertinent pour l'évaluation de la fluidité graphique chez l'enfant d'âge scolaire.

**Mots-clés.** Neuropsychologie clinique, fonctions exécutives, fluidité graphique, technologie, validation

## **Abstract**

Rapid technological advancements have revolutionized various industries, yet the field of neuropsychology is just starting to embrace these changes. The integration of computerized tests into practice has been insufficient, despite the numerous advantages they offer. These advantages include enhanced test fidelity and validity, time and cost savings, and improved accessibility to neuropsychological services. In light of these considerations, we developed an electronic version of the Five-Point Test in the form of a mobile application called AppFLO. The aim of this project was to conduct a preliminary validation of the AppFLO test with children aged 6 to 12. A battery of tests and questionnaires was administered to 30 neurotypical children to evaluate the fidelity and validity of the tool. Results revealed that AppFLO performance improves with age, school level, and intellectual ability, and is independent of gender and parent's educational level. In addition, our findings showed that the test has good test-retest reliability. Moderate to strong correlations between the AppFLO, the Five-Point Test and other tests measuring executive functions suggest good construct validity. In conclusion, this preliminary validation study indicates that the AppFLO test holds promise as an effective tool for assessing design fluency in school-age children.

**Mots-clés.** Clinical neuropsychology, executive function, design fluency, technology, validation

## Liste des tableaux

Tableau 1. <i>Définition des différentes variables de AppFLO</i> .....	8
Tableau 2. <i>Caractéristiques sociodémographiques de l'échantillon</i> .....	15
Tableau 3. <i>Moyenne et écart-type des scores au AppFLO en fonction de l'âge</i> .....	16
Tableau 4. <i>Coefficients de corrélation de Pearson entre les scores au AppFLO et l'âge, le niveau scolaire et le QI estimé (Similitudes et Matrices)</i> .....	17
Tableau 5. <i>ANOVAs de l'effet du sexe et du niveau d'éducation de la mère et du père sur la performance au AppFLO</i> .....	17
Tableau 6. <i>Comparaison des performances au AppFLO et au FPT</i> .....	19
Tableau 7. <i>Coefficients de corrélation de Pearson entre les scores principaux au AppFLO et les scores en Fluidité verbale, à la Tour de Londres, au test d'Empan spatial et au Stroop</i> .....	20

## Liste des figures

Figure 1. <i>Visuel et dimensions standardisées du AppFLO</i> .....	9
---	---

## Liste des abréviations

$\alpha_{\text{critique}}$  : Alpha critique

ANOVA : Analyse de variance

BRIEF : Inventaire de comportements reliés aux fonctions exécutives

C-DF : *Computerized Test of Design Fluency*

CEG : Score composite exécutif global

É.-T. : Écart-type

FE : Fonctions exécutives

FG : Fluidité graphique

FPT : Five-Point Test

ICC : Coefficient de corrélation intraclasse

IM : Indice de métacognition

IRC : Indice de régulation comportementale

M : Moyenne

Nb : Nombre

$p$  : Seuil de signification

QI : Quotient intellectuel

TDA/H : trouble de déficit de l'attention/hyperactivité

WISC-V : Test d'intelligence de Wechsler, 5<sup>e</sup> édition

## Introduction

Dans les dernières années, nous avons assisté à une avancée fulgurante de l'utilisation des technologies dans les différents domaines médicaux (Miller et Barr, 2017). Cependant, la profession de neuropsychologue ne semble pas s'être adaptée à ces changements. Actuellement, la majorité des tests utilisés par les neuropsychologues demeure en version papier (Rabin et al., 2016; Rabin et al., 2014). Il s'agit principalement des mêmes outils qui étaient utilisés au début des années 2000 (Rabin et al., 2016), et qui ont, pour la plupart, été développés au cours du 20<sup>e</sup> siècle (Bilder et Reise, 2019). Bien que ces outils soient utiles tant en recherche que sur le plan clinique, ils ne sont pas représentatifs des avancées majeures observées dans les dernières années quant aux théories en psychométrie, aux technologies et à la compréhension des relations entre le cerveau et le comportement (Bilder et Reise, 2019). Selon Miller and Barr (2017), il sera crucial d'intégrer davantage la technologie au sein de la neuropsychologie afin de préserver la pertinence de cette discipline. Parmi les nombreux avantages à utiliser les outils numériques se trouvent une diminution des erreurs d'administration et de cotation, une économie de temps ainsi qu'une amélioration globale de l'efficacité clinique (Miller et Barr, 2017). De plus, les mesures technologiques seraient plus engageantes et motivantes que les mesures traditionnelles (Schultheis et Doiron, 2017).

Face à cette évolution technologique rapide, quelques batteries de tests en format numérique permettent d'évaluer, entre autres, les fonctions exécutives, tels que la NIH EXAMINER et la batterie CANTAB (*Cambridge Neurological Test Automated Battery*). Ces batteries demeurent peu utilisées en clinique au Québec en raison du peu d'études de validation et l'absence de normes auprès de la population québécoise. Ainsi, un manque d'outils numériques sévit dans le contexte clinique. La présente étude vise à prévalider un nouveau test numérique d'évaluation des fonctions exécutives (FE) afin de pallier à ce manque.

Les FE sont un ensemble de processus cognitifs de haut niveau mis en branle lors d'activités cognitives complexes ou lors de situations inhabituelles et nouvelles (Lussier et al., 2017). Elles jouent un rôle crucial dans la réussite scolaire ainsi que dans le fonctionnement comportemental et social des enfants (Cortés Pascual et al., 2019; Holmes et al., 2016; Riggs et al., 2004; Wang et Liu, 2021; Wang et Zhou, 2019). Il y aurait trois composantes de base des FE, soit l'inhibition, la mémoire de travail et la flexibilité cognitive, desquelles découleraient des fonctions de plus haut niveau, telles que la planification et l'organisation (Diamond, 2013; Lehto et al., 2003; Miyake et al., 2000). La fluidité graphique (FG) est une autre de ces fonctions. Il importe de distinguer la tâche de FG de la fonction cognitive. En effet, la tâche de FG fait référence au test qui évalue l'habileté à initier et maintenir une production de différentes figures abstraites tout en utilisant des stratégies dans le but d'éviter la répétition et maximiser la production de réponses (Cattelani et al., 2011; Tucha et al., 2012). Pour sa part, la FG en tant que fonction exécutive fait référence à l'ensemble des fonctions cognitives impliquées dans l'exécution de cette tâche, comme la planification, l'initiation, la résolution de problème, la flexibilité cognitive, l'inhibition, la mémoire de travail, l'attention visuelle, les habiletés visuospatiales et constructives ainsi que la motricité (Delis et al., 2001; Ruff, 1996; Suchy et al., 2010).

La performance en FG serait indépendante du genre (Albert et al., 2009; Klenberg et al., 2001; Matias-Guiu et al., 2021; Regard et al., 1982; Van der Elst et al., 2011), mais s'améliorerait avec l'âge, le niveau scolaire de l'enfant et le niveau d'éducation des parents (Klenberg et al., 2001; Levin et al., 1991; Regard et al., 1982; Van der Elst et al., 2011). Différentes études se sont aussi intéressées au lien entre la performance en fluidité graphique et le fonctionnement intellectuel. Les résultats sont inconstants : plusieurs études observent un lien entre ces deux fonctions (Arffa, 2007; Goebel et al., 2009; Lee et al., 1997; Regard et al., 1982; Vik et Ruff, 1988), mais d'autres n'obtiennent pas de corrélations (Hurks, 2013; Tucha et al., 2012). D'autre part, on observerait des corrélations entre la fluidité graphique et les

autres fonctions exécutives telles que la planification, la fluidité verbale, la mémoire de travail, l'inhibition et la flexibilité cognitive (Goebel et al., 2009; Stievano et Scalisi, 2016; Tucha et al., 2012) En effet, des régions anatomiques communes, soit le cortex dorsolatéral préfrontal, seraient impliquées pour l'ensemble de ces fonctions, bien que chacune d'entre elles dépendrait également de régions cérébrales plus spécifiques (Ghanavati et al., 2019; Nejati et al., 2018).

Il existe différents tests de FG en modalité papier qui peuvent être utilisés en neuropsychologie dans le but d'évaluer les FE (Baldo et al., 2001; Goebel et al., 2009). Il s'agit de mesures sensibles au fonctionnement ainsi qu'aux lésions du lobe frontal (Fernandez et al., 2008; Lee et al., 1997). Chez les enfants, certains troubles neurodéveloppementaux, comme la dyslexie et le trouble de déficit de l'attention/hyperactivité (TDA/H), semblent influencer les résultats en FG (Gauthier et al., 2016; Reiter et al., 2005; Robinson et Tripp, 2013; Vélez-van-Meerbeke et al., 2013).

Le Five-Point Test (FPT) est un test de FG qui se présente sur une feuille 8½ x 11 pouces comprenant 40 rectangles disposés en matrice de 8 lignes x 5 colonnes. Chacun des rectangles contient cinq points, qui sont disposés comme sur la face d'un dé. L'objectif de la tâche est de produire le plus de dessins différents possible en reliant les points avec des lignes droites dans un temps variant de 2 à 5 minutes, selon les versions (Cattelani et al., 2011; Regard et al., 1982; Tucha et al., 2012). Différents scores peuvent être considérés lors de l'analyse des résultats au FPT : le nombre total de dessins, le nombre et le pourcentage de répétitions, le nombre et le pourcentage de dessins corrects (dessins qui ne sont pas des répétitions ou des bris de règles) et le nombre de stratégies (Regard et al., 1982; Stievano et Scalisi, 2016). Des stratégies numériques (ajout ou retrait d'une ligne au dessin suivant) ainsi que des stratégies rotatives et miroirs peuvent être utilisées (Regard et al., 1982). Selon Lezak (1995), une utilisation adéquate de stratégies au FPT permettrait de maximiser la production de dessins tout en permettant une réduction du nombre de répétitions, puisqu'une série de dessins différents pourraient être produits à partir

d'une seule idée. Des études de Stievano and Scalisi (2016) et de Goebel et al. (2009) permettent de confirmer que les enfants qui présentaient un pourcentage plus élevé de stratégies produisaient plus de dessins total et corrects. Concernant la relation entre le pourcentage de stratégies et le pourcentage de répétitions, les résultats sont variables. Goebel et al. (2009) observent une corrélation négative entre les mesures, ce qui corrobore l'hypothèse de Lezak (1995), alors que Stievano and Scalisi (2016) notent une absence de corrélation. D'autre part, les résultats de Stievano and Scalisi (2016) indiquent que plus un enfant effectuait de dessins total, plus il effectuait un pourcentage élevé de répétitions. En fait, plus un enfant produit de dessins, plus le nombre de nouveaux graphiques possible diminue, ce qui demande un plus grand contrôle pour éviter les répétitions (Gauthier et al., 2016) et pourrait expliquer les résultats obtenus dans cette étude.

Le FPT présente une excellente fidélité interjuge et une bonne fidélité test-retest (Fernandez et al., 2008; Skalla, 2012; Tucha et al., 2012). Il présente toutefois un certain effet de pratique, puisque le nombre de dessins corrects produits tend à augmenter lors de la deuxième passation (Fernandez et al., 2008; Tucha et al., 2012). Par ailleurs, le nombre de dessins corrects produit au FPT corrèle à la performance à d'autres tests neuropsychologiques mesurant des construits similaires, tels que la Tour de Londres (temps d'initiation), le test d'interférence mot-couleur (condition interférence), le test de fluidité verbale (nombre de mots produits à la condition de lettres et sémantique) et les blocs de Corsi (Stievano et Scalisi, 2016; Tucha et al., 2012). De plus, le pourcentage de répétitions corrèle à l'empan de chiffres à l'envers de la Wechsler Memory Scale Revised Edition (WMS-R) et à la performance aux blocs de Corsi (Goebel et al., 2009; Stievano et Scalisi, 2016), indiquant que cette mesure converge avec les mesures de mémoire de travail. La corrélation entre le FPT et des questionnaires sur le fonctionnement exécutif n'a jamais été explorée pour examiner la relation entre les mesures formelles de FE et les mesures par questionnaires. Cependant, selon les résultats d'une méta-analyse (Toplak et al., 2013), les résultats aux questionnaires

évaluant les FE sur le plan comportemental sont généralement très faiblement associés aux résultats obtenus dans des tests neuropsychologiques des FE basés sur la performance.

Malgré ses bonnes qualités psychométriques, le FPT présente certaines limites. Par exemple, la correction peut être longue et engendrer des erreurs, l'analyse de la performance se complexifie à mesure que les intervalles de temps considérés diminuent, et l'aspect moteur est difficilement pris en considération. La technologie permettrait de contourner ces limites. Dans cette optique, Woods et al. (2016) ont développé un test de FG sur ordinateur, le *Computerized Test of Design Fluency (C-DF)*. Dans ce test, cinq points disposés aléatoirement sont présentés à l'écran. L'objectif du test est de produire le plus de dessins différents en 90 secondes en connectant les cinq points avec quatre lignes continues. Une souris d'ordinateur est utilisée pour effectuer les dessins. Dans le C-DF, les dessins déjà produits ne peuvent être vus, contrairement aux tests de FG papier. Selon Woods et al. (2016), cette façon de fonctionner est plus exigeante au niveau mnésique qu'un test de FG en version papier dans lequel les dessins précédents demeurent visibles. En effet, les créateurs du test informatisé ont observé un pourcentage plus élevé de répétitions avec le C-DF que les versions papier. Les participants produisaient également moins de dessins corrects au C-DF. Les auteurs expliquent cette différence par le fait qu'il est peu commun de dessiner avec une souris (Woods et al., 2016).

Dans le but de contourner les limites de la souris et répondre au besoin criant d'outils numériques en neuropsychologie clinique, nous avons développé un test sur tablette numérique évaluant la fluidité graphique qui est inspiré du FPT, le test AppFLO. L'objectif de la présente étude est de prévalider le test AppFLO chez des enfants de 6 à 12 ans. Plus précisément, il s'agit d'une étude pilote visant à vérifier la faisabilité d'une validation plus large de l'outil. Nous observerons également de façon préliminaire la fidélité test-retest, l'équivalence entre la version papier et la version numérique du test ainsi que la validité de construit de l'outil. Si la prévalidation confirme la faisabilité de l'étude et qu'elle procure des qualités

psychométriques adéquates de l’outil, une validation à plus large échelle et la normalisation pourront être envisagées. Cela permettra d’offrir à la communauté de cliniciens un outil technologique utile, rapide et efficace, et ainsi de contribuer à moderniser les pratiques en neuropsychologie clinique.

## **Méthodologie**

### **Participants**

Un total de 48 enfants âgés de 6 à 12 ans ont participé à l’étude (25 filles, 23 garçons). Pour être inclus dans les analyses, les participants devaient parler le français et ne devaient pas avoir de diagnostic de troubles neurologiques, de santé mentale ou neurodéveloppementaux (c.-à-d. de trouble de la communication ; de trouble du spectre de l’autisme ; de TDA/H ; de trouble d’apprentissage ; de retard intellectuel ou de trouble de tics - voir la cinquième édition du *Manuel diagnostique et statistique des troubles mentaux* (American Psychiatric Association, 2013)). Un questionnaire sociodémographique rempli par le parent a permis de vérifier les critères d’exclusion.

### **Procédure**

Le recrutement a été effectué dans des écoles de Montréal et de Québec à l’aide d’un formulaire de sollicitation qui a été distribué dans les classes ainsi qu’un formulaire numérique qui a été transmis aux parents par courriel de la part des directions d’écoles. Les rencontres de passation de tests ont eu lieu dans des locaux de recherche de l’Université de Montréal ou dans une clinique privée à Québec. Au début de la première séance d’évaluation, les parents ont signé le formulaire de consentement qu’ils ont préalablement reçu par courriel. Lors de la première séance, d’une durée d’environ deux heures, une batterie de tests neuropsychologiques a été administrée, toujours dans le même ordre : AppFLO, Similitudes, Empan spatial, Purdue Pegboard, Interférence couleur-mots, Matrices, Fluidité verbale, Tour de Londres, FPT, Estimation Line et Beery VMI condition motrice. AppFLO était le premier test administré afin d’éviter un effet de pratique à la suite du FPT. Après avoir complété AppFLO, l’enfant

devait indiquer son appréciation de la tâche à l'aide d'une échelle de Likert imagée à cinq points. Pendant ce temps, les parents remplissaient des questionnaires concernant leur enfant: un questionnaire sociodémographique et l'Inventaire de comportements reliés aux fonctions exécutives (BRIEF) (Gioia et al., 2000). Une seconde séance d'environ 45 minutes avait lieu, si possible, 14 à 28 jours après la première (M = 23,71, É.-T. = 12,17) dans le but d'évaluer la fidélité du AppFLO. Les tests administrés lors de cette séance étaient, dans l'ordre : AppFlo, Tempo Test Reneken, Numeracy Screener, Path Span, Continuous Attention Test. Puisqu'il s'agit d'une étude plus large, certains tests ont été administrés pour d'autres objectifs (voir section Matériel pour la description des outils utilisés dans la présente étude). Par ailleurs, tous les tests des participants ont été double-corrigés à la main afin de s'assurer que les résultats étaient les mêmes que ceux qui seront obtenus dans la version finale de l'application, puisqu'il s'agissait d'une version alpha du AppFLO. Les participants ont reçu une compensation financière de 15\$. Le projet a été approuvé par le Comité d'éthique à la recherche en éducation et psychologie de l'Université de Montréal (Certificat no CEREP-20-146-P) et le Comité de la recherche du Centre de services scolaire de Montréal (Demande d'expérimentation 2020-2021).

## **Matériel**

### **Questionnaire**

*Behavioral Rating Inventory of Executive Function (BRIEF) - Parent Form* (Gioia et al., 2000).

Ce questionnaire mesure les comportements liés aux fonctions exécutives chez les jeunes âgés de 5 à 18 ans. Il comprend 86 items regroupés en huit échelles de fonctions exécutives : inhibition, alternance/flexibilité, contrôle émotionnel, initiation, mémoire de travail, planification, organisation et contrôle. Les trois premières échelles forment un indice de régulation comportementale (IRC). Les suivantes forment un indice de métacognition (IM). Puis, ces deux indices forment un score composite exécutif global (CEG). Ce questionnaire possède une bonne fidélité (cohérence interne entre 0,80 et 0,98

et fidélité test-retest entre 0,76 et 0,92) (Strauss et al., 2006). Les scores bruts des différentes échelles et indices seront utilisés pour vérifier la relation avec le AppFLO.

### Tests neuropsychologiques

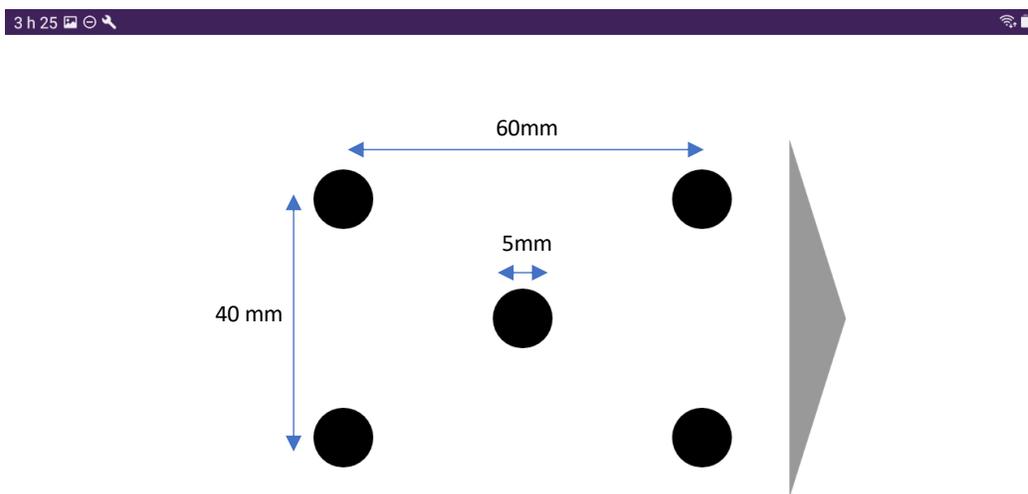
**Le test AppFLO.** Le AppFLO est un test de fluidité graphique qui s'administre sur tablette électronique. Cinq points sont présentés à l'écran. Les points ont la même taille et sont disposés de manière standardisée, peu importe la tablette utilisée pour l'évaluation (voir Figure 1). La tâche consiste à faire un dessin à l'aide du doigt et d'appuyer sur une flèche située à droite de l'écran lorsque celui-ci est complété afin de faire un nouveau dessin. L'objectif est de produire le plus de dessins différents possible en 5 minutes, et ce, en utilisant des lignes droites (traits orange d'une épaisseur de 5 pixels) qui commencent à un point et se terminent à un autre point. Les variables collectées automatiquement par l'application qui sont prises en compte pour les analyses de la présente étude sont les suivantes : Nombre total de dessins, Nombre total de répétitions, Nombre de répétitions adjacentes, Nombre de répétitions non adjacentes, Nombre de stratégies, Nombre de dessins invalides, Nombre de dessins valides, Nombre de dessins corrects, Pourcentage de répétitions, Pourcentage de dessins invalides et Pourcentage de dessins corrects. Pour la validité de construit, cinq variables principales sont considérées : Nombre total de dessins, Nombre de dessins corrects, Nombre de stratégies, Nombre total de répétitions et Pourcentage de répétitions. Il s'agit des mesures les plus souvent étudiées en recherche ou utilisées en clinique.

**Tableau 1**

*Définition des différentes variables de AppFLO*

Nombre total de dessins	Nombre de dessins effectués, incluant les dessins répétés et invalides.
Nombre total de répétitions	Nombre de dessins identiques à un dessin précédent, adjacents ou non.
Nombre de répétitions adjacentes	Nombre de dessins identiques au dessin précédent.
Nombre de répétitions non adjacentes	Nombre de dessins identiques à un dessin précédent, séparé de celui-ci par un ou d'autres dessins.
Nombre total de stratégies	Nombre de dessins qui suggèrent l'utilisation d'une stratégie spatiale (rotation ou miroir) ou numérique (ajout ou retrait d'une ligne) appliquée au dessin précédent.

Nombre de dessins invalides	Nombre de dessins qui ne respectent pas les règles de relier les points avec des lignes droites et qui vont d'un point à un autre point.
Nombre de dessins valides	Nombre de dessins, incluant les répétitions, qui ne sont pas des dessins invalides.
Nombre de dessins corrects	Nombre de dessins qui ne sont pas des répétitions ou des dessins invalides.
Pourcentage de répétitions	Nombre total de répétitions / Nombre total de dessins * 100
Pourcentage de dessins invalides	Nombre de dessins invalides / Nombre total de dessins * 100
Pourcentage de dessins corrects	Nombre de dessins corrects / Nombre total de dessins * 100



**Figure 1.** Visuel et dimensions standardisées du AppFLO

**Matrices et Similitudes** (Wechsler, 2015). Les tests Matrices et Similitudes de l'Échelle d'intelligence de Wechsler pour enfants – 5e édition : version pour francophones du Canada (WISC-V) ont été administrés afin d'obtenir un estimé du quotient intellectuel (QI) verbal et non verbal, respectivement. Ces tests sont ceux qui corréleront le mieux avec le QI global (respectivement 0,69 et 0,73 pour la WISC-V). Dans le test Matrices, le participant doit choisir l'item qui complète une matrice incomplète parmi cinq choix (raisonnement non verbal). Durant le test Similitudes, il est demandé au participant de définir le lien conceptuel qui unit deux mots (raisonnement verbal). La fidélité test-retest de

ces tests est élevée (entre 0,80 et 0,89) (Wechsler, 2015). L'ajout de ces tests au protocole permettra de vérifier la relation entre le QI estimé (scores bruts aux tests de Matrices et Similitudes) et la performance au AppFLO.

***Empan Spatial*** (Wechsler et al., 2004). Ce test de l'Échelle d'intelligence de Wechsler pour enfants – 4e édition : version intégrée mesure la mémoire de travail visuelle et comprend deux conditions. Dans la première condition, qui mesure l'encodage, l'administrateur pointe des blocs dans un ordre précis et le participant doit reproduire la séquence dans le même ordre. Dans la seconde condition, qui mesure la mémoire de travail, le participant doit reproduire la séquence dans l'ordre inverse. La cohérence interne et la fidélité test-retest sont adéquates pour ce test (entre 0,70 et 0,79) (Strauss et al., 2006). Les scores bruts de chaque condition seront utilisés pour évaluer la validité de construit.

***Interférence couleur-mot*** (Delis et al., 2001). Ce test mesure l'inhibition et la flexibilité cognitive. Il comprend quatre conditions. Dans la condition 1, le participant doit dénommer des couleurs. Dans la condition 2, il doit lire des noms de couleurs. Dans les conditions suivantes, les noms de couleurs sont imprimés dans une couleur d'encre différente. Dans la condition 3, le participant doit nommer la couleur de l'encre et non lire le mot (inhibition). Dans la condition 4, il doit alterner entre nommer la couleur de l'encre et lire le mot en fonction d'un indice visuel (inhibition et flexibilité cognitive). La cohérence interne pour ce test est adéquate (0,70 à 0,79) et la fidélité test-retest varie de marginale à adéquate (entre 0,60 et 0,79) (Strauss et al., 2006). Les scores utilisés pour la validité convergente seront le score de temps et le nombre d'erreurs des conditions 3 et 4.

***Fluidité verbale*** (Delis et al., 2001). Il s'agit d'un test qui mesure la fluence verbale et la flexibilité cognitive. Le test comprend trois conditions, qui comprennent plusieurs essais de 60 secondes. Durant la tâche, le participant doit générer le plus de mots possible qui débutent par une lettre précise (condition de lettres), une catégorie précise (condition sémantique) et en alternant entre deux catégories (condition

alternance). La fidélité test-retest pour ce test varie d'adéquate à élevée (entre 0,70 et 0,89) et la cohérence interne varie de marginale à élevée (entre 0,60 et 0,89) (Strauss et al., 2006). Les scores d'intérêts qui seront utilisés pour la validité convergente sont le nombre total de mots pour les différentes conditions ainsi que le nombre d'alternances correctes pour la troisième condition.

**Tour de Londres-Drexel (TOLDX)** (Culbertson et al., 1997). Ce test mesure la planification et la résolution de problème. Le participant doit reproduire les modèles d'une tour à l'aide de billes de couleur sur des tiges de bois en respectant certaines règles. Le test présente une bonne fidélité test-retest pour le nombre total de mouvements (0,81) (Culbertson et Zillmer, 1998). Le nombre total de mouvements, qui évalue la planification, et le temps d'initiation, soit le délai avant le premier mouvement ou le temps pris pour la planifier, seront utilisés pour la validité convergente.

**Five-Point Test** (Regard et al., 1982). Se référer à la présentation dans l'introduction. Rappelons tout de même que le FPT présente une excellente fidélité interjuge, une bonne fidélité test-retest et une bonne validité de construit (Fernandez et al., 2008; Goebel et al., 2009; Skalla, 2012; Stievano et Scalisi, 2016; Tucha et al., 2012).

### **Analyse des données et prédictions**

Les analyses ont été effectuées avec le logiciel IBM SPSS Statistics 28. Le seuil de signification utilisé était de ,05 ( $\alpha_{critique} = ,05$ ). Étant donné le petit échantillon de participants, les valeurs de  $p$  qui s'approchaient du seuil de signification (entre ,051 et ,065) ont également été indiquées dans les tableaux. Selon le théorème de limite central (Kwak et Kim, 2017), la taille de l'échantillon de la présente étude ( $N=30$ ) permet de présupposer la normalité de la distribution des différentes variables pour les analyses.

Afin de vérifier la faisabilité de l'étude, le pourcentage d'enfants ayant complété le protocole de recherche (taux d'attrition), le pourcentage d'enfant ayant effectué la tâche du AppFLO jusqu'à la fin et un score d'appréciation de la tâche à partir d'une échelle de Likert à cinq points ont été calculés.

Des corrélations de Pearson ont été calculées entre l'âge, le niveau scolaire ainsi que la performance aux tests de Matrices et Similitudes (QI estimé) et la performance au AppFLO. Il était attendu que l'âge, le niveau scolaire et le QI estimé soient corrélés positivement aux différentes mesures du AppFLO. Des ANOVAs ont également été effectuées afin de vérifier si la performance au AppFLO était différente selon le sexe de l'enfant et le niveau d'éducation des parents. Il était attendu qu'il n'y ait pas de différence significative selon le sexe, mais que la performance des enfants dont les parents avaient un niveau d'éducation plus élevé soit meilleure.

La fidélité test-retest du AppFLO a été mesurée en calculant la corrélation intraclasse entre les deux temps de mesures pour les différents scores. Il était attendu que les scores aux deux temps de mesure corrélaient modérément entre eux. Des tests-t à échantillon appariés ont été conduits afin de vérifier si la performance s'améliorait lors de la deuxième passation. Il était attendu que les enfants produisent davantage de dessins corrects au second temps de mesure.

Afin de mesurer la validité de construit, les scores principaux du AppFLO (Nombre total de dessins, Nombre de dessins corrects, Nombre de stratégies, Nombre total de répétitions et Pourcentage de répétitions) ont été corrélés (Pearson) aux scores des tests Interférence couleur-mot, Fluidité verbale, Empan spatial, Tour de Londres ainsi qu'aux scores au BRIEF. Il était attendu que le Nombre de dessins corrects au AppFLO corrélaient positivement au nombre total de mots dans les conditions sémantique et de lettres en Fluidité verbale, au temps d'initiation à la Tour de Londres, aux scores de l'Empan spatial à l'endroit et à l'envers et à la vitesse à la condition 3 du Stroop. De plus, il était attendu que le Nombre de stratégies soit corrélé positivement au nombre total de mouvements à la Tour de Londres et que le Pourcentage de répétitions soit corrélé négativement aux scores des deux conditions d'Empan Spatial. Il était attendu qu'il n'y ait pas de relation entre la performance au AppFLO et les résultats au BRIEF. Les scores principaux du AppFLO ont également été corrélés entre eux afin d'évaluer la validité de construit

intra-test de l'outil. Une corrélation positive était attendue entre le Nombre de stratégies et le Nombre de dessins total et correct. Une corrélation positive était également attendue entre le Nombre total de dessins et le Nombre total de répétitions. Les corrélations entre le Nombre de stratégies ainsi que le Nombre total et le Pourcentage de répétitions étaient exploratoires considérant les résultats variables dans la littérature actuelle.

L'équivalence entre la version papier et la version numérique a été vérifiée grâce à une corrélation de Pearson entre le AppFLO et le FPT pour les différents scores. Une corrélation était attendue entre les deux versions du test. Des tests-t à échantillon appariés ont ensuite été conduits afin de vérifier si les mesures différaient selon la modalité du test. Un nombre plus élevé de répétitions était attendu pour la version numérique.

## **Résultats**

### **Description de l'échantillon**

Sur un total de 48 participants potentiels, 18 participants ont été exclus. Dix d'entre eux avaient reçu un diagnostic officiel d'un trouble neurodéveloppemental, sept ont rencontré des problèmes lors de la passation ou l'enregistrement du test AppFLO sur la tablette, et un présentait des scores à deux écarts-types sous la moyenne au AppFLO. Ainsi, l'échantillon final est composé de 30 enfants (19 filles, 11 garçons) âgés de 6 à 12 ans ( $M = 9,39$ ,  $É.-T.=1,52$ ). Pour les études pilotes de ce type, des groupes de 12 à 30 participants sont généralement recommandés (Browne, 1995; Julious, 2005). Les enfants étaient majoritairement droitiers (90%) et caucasiens (76,7%). La majorité des parents avaient fait des études universitaires (baccalauréat, maîtrise et/ou doctorat), tant pour la mère (66,7%) que pour le père (63,3%). Les caractéristiques sociodémographiques de l'échantillon final sont présentées dans le tableau 2. Pour les analyses concernant la fidélité test-retest, seuls 25 participants ont été inclus, car deux des 30 participants

n'étaient pas disponibles pour une deuxième rencontre et certains problèmes ont été rencontrés avec trois d'entre eux lors de la passation ou l'enregistrement du test sur la tablette à la deuxième rencontre.

### **Faisabilité de l'étude**

D'abord, la totalité des participants (100%) a complété l'ensemble des tâches prévues à la première rencontre d'évaluation. De plus, l'ensemble des participants (100%) a complété la passation du AppFLO d'une durée de 5 minutes. Le taux d'attrition pour la deuxième rencontre d'évaluation était de 6,7%. Finalement, le score moyen du niveau d'appréciation des enfants face à la tâche du AppFLO est de  $4,21 \pm 0,56$ . Sur une échelle de Likert à cinq points, un enfant n'a pas répondu, trois enfants ont estimé leur appréciation à 3/5, dix-neuf enfants l'ont estimé à 4/5 et huit enfants l'ont estimé à 5/5.

### **Performance, facteurs sociodémographiques et QI**

Le tableau 3 présente les résultats des enfants au AppFLO en fonction de leur âge. L'âge et le niveau scolaire étaient corrélés positivement et modérément au Nombre de dessins corrects et au Nombre de stratégies au AppFLO (Tableau 4). De plus, une corrélation négative et modérée a été observée entre le niveau scolaire et le Pourcentage de répétitions. Une corrélation positive et modérée a également été observée entre le niveau scolaire et le Pourcentage de dessins corrects (Tableau 4). Des corrélations positives quasi significatives, mais de degré modéré, ont été observées entre l'âge et le Nombre total de dessins ( $p = ,052$ ), le Nombre de dessins valides ( $p = ,052$ ) et le Pourcentage de dessins corrects ( $p = ,060$ ). Des corrélations positives quasi significatives, mais de degré modéré, ont aussi été observées entre le niveau scolaire, le Nombre total de dessins ( $p = ,062$ ) et le Nombre de dessins valides ( $p = ,057$ ). Par ailleurs, des corrélations de modérées à fortes ont été observées entre le QI estimé et certains scores au AppFLO (Tableau 4). Il n'y avait pas de différences significatives entre la performance des filles et des garçons, et ce, pour tous les scores au AppFLO (Tableau 5). De même, les scores aux AppFLO n'étaient pas différents selon le niveau d'éducation des parents (Tableau 5).

**Tableau 2***Caractéristiques sociodémographiques de l'échantillon*

	N	%
Sexe		
Féminin	19	63,3
Masculin	11	36,7
Âge		
6	1	3,3
7	7	23,3
8	5	16,7
9	4	13,3
10	6	20,0
11	6	20,0
12	1	3,3
Année scolaire		
1 <sup>e</sup>	4	13,3
2 <sup>e</sup>	7	23,3
3 <sup>e</sup>	4	13,3
4 <sup>e</sup>	6	20,0
5 <sup>e</sup>	8	26,7
6 <sup>e</sup>	1	3,3
Revenu familial brut annuel		
20 000 à 39 000	1	3,3
40 000 à 59 000	2	6,7
60 000 à 79 000	3	10,0
100 000 et plus	23	76,7
Non disponible	1	3,3
Diplôme le plus élevé de la mère		
Baccalauréat	11	36,7
Maîtrise et/ou doctorat	9	30,0
Autre	8	26,7
Non disponible	2	6,7
Diplôme le plus élevé du père		
Baccalauréat	12	40,0
Maîtrise et/ou doctorat	7	23,3
Autre	6	20,0
Non disponible	5	16,7
Région		
Montréal	18	60,0
Québec	12	40,0

**Tableau 3***Moyenne et écart-type des scores au AppFLO en fonction de l'âge*

		6	7	8	9	10	11	12
		n = 1	n = 7	n = 5	n = 4	n = 6	n = 6	n = 1
Nb total de dessins	M	16,00	29,14	34,60	41,75	33,33	44,17	21,00
	É.-T.	-	10,22	10,46	11,59	11,61	15,69	-
Nb total de répétitions	M	4,00	6,43	5,00	6,25	3,17	9,00	0,00
	É.-T.	-	5,77	3,16	5,32	3,06	6,99	-
Nb de répétitions adjacentes	M	1,00	0,57	0,40	0,25	0,67	0,67	0,00
	É.-T.	-	1,14	0,55	0,50	1,03	1,211	-
Nb de répétitions non adjacentes	M	3,00	5,86	4,60	6,00	2,50	8,33	0,00
	É.-T.	-	5,05	2,70	4,83	2,26	6,47	-
Nb de stratégies	M	3,00	5,57	5,40	10,25	5,33	16,00	1,00
	É.-T.	-	2,76	3,21	3,86	4,68	11,87	-
Nb de dessins invalides	M	2,00	0,29	1,20	0,50	1,17	0,83	0,00
	É.-T.	-	0,49	2,17	0,58	2,40	1,60	-
Nb de dessins valides	M	14,00	28,86	33,40	41,25	32,17	43,33	21,00
	É.-T.	-	10,19	9,07	11,24	11,30	15,48	-
Nb de dessins corrects	M	10,00	22,43	28,40	35,00	29,00	34,67	21,00
	É.-T.	-	9,18	6,58	6,38	9,38	12,85	-
Pourcentage de répétitions	M	25,00	20,76	13,44	13,60	8,42	18,17	0,00
	É.-T.	-	17,04	6,57	7,60	6,64	14,34	-
Pourcentage de dessins invalides	M	12,50	0,94	2,95	1,10	3,09	1,71	0,00
	É.-T.	-	1,60	4,45	1,33	6,50	3,33	-
Pourcentage de dessins corrects	M	62,50	78,31	83,60	85,30	88,49	81,07	100,00
	É.-T.	-	17,72	8,06	8,41	10,55	17,35	-

**Tableau 4**

Coefficients de corrélation de Pearson entre les scores au AppFLO et l'âge, le niveau scolaire et le QI estimé (Similitudes et Matrices)

	Âge	Niveau scolaire	Similitudes	Matrices
	<i>r</i>			
Nb total de dessins	,30 <sup>††</sup>	,29 <sup>†</sup>	,35 <sup>†</sup>	,23
Nb total de répétitions	< ,01	-,05	-,03	-,13
Nb de répétitions adjacentes	,02	-,03	-,26	-,35 <sup>†</sup>
Nb de répétitions non adjacentes	< -,01	-,05	,02	-,07
Nb de stratégies	,33*	,33*	,34 <sup>†</sup>	,28
Nb de dessins invalides	,05	-,02	-,09	-,39*
Nb de dessins valides	,30 <sup>††</sup>	,29 <sup>†</sup>	,37*	,28
Nb de dessins corrects	,38*	,40*	,48**	,41*
Pourcentage de répétitions	-,28	-,32*	-,30	-,36*
Pourcentage de dessins invalides	-,09	-,13	-,20	-,56**
Pourcentage de dessins corrects	,29 <sup>†</sup>	,35*	,34 <sup>†</sup>	,50**

Notes. † = ,055 < p < ,065, †† = ,050 < p < ,055, \* = p < ,050, \*\* = p < ,010.

**Tableau 5**

ANOVAs de l'effet du sexe et du niveau d'éducation de la mère et du père sur la performance au AppFLO

	Sexe	Niveau d'éducation de la mère	Niveau d'éducation du père
	<i>F</i> (1,28)	<i>F</i> (2,25)	<i>F</i> (2,22)
Nb total de dessins	0,005	1,524	1,207
Nb total de répétitions	1,835	0,138	0,182
Nb de répétitions adjacentes	1,482	0,800	1,897
Nb de répétitions non adjacentes	1,559	0,064	0,191
Nb de stratégies	0,284	1,181	0,983
Nb de dessins invalides	0,037	0,123	0,854
Nb de dessins valides	0,009	1,541	1,356
Nb de dessins corrects	0,674	2,377	2,591
Pourcentage de répétitions	1,125	0,456	1,035
Pourcentage de dessins invalides	0,003	0,098	1,630
Pourcentage de dessins corrects	1,021	0,569	1,486

Notes. \* = p < ,05, \*\* = p < ,01.

### Fidélité test-retest

Des degrés d'accord élevés entre les deux temps de mesure ont été obtenus pour le Nombre de stratégies (ICC = ,79, p < ,001), le Nombre de dessins corrects (ICC = ,75, p < ,001), le Pourcentage de

répétitions ( $ICC = ,82, p < ,001$ ) et le Pourcentage de dessins corrects ( $ICC = ,86, p < ,001$ ). De plus, des degrés d'accord modérés entre les deux temps de mesures ont été obtenus pour le Nombre total de dessins ( $ICC = ,69, p = ,002$ ), le Nombre total de répétitions ( $ICC = ,71, p = ,002$ ), le Nombre de répétitions adjacentes ( $ICC = ,73, p = ,001$ ), le Nombre de répétitions non adjacentes ( $ICC = ,56, p = ,025$ ) et le Nombre de dessins valides ( $ICC = ,70, p = ,001$ ). Cependant, il n'y avait pas de corrélations significatives entre les deux temps de mesure pour le Nombre de dessins invalides ( $ICC = ,41, p = ,108$ ) et le Pourcentage de dessins invalides ( $ICC = ,37, p = ,138$ ).

Les enfants produisaient significativement plus de dessins total et corrects à la deuxième passation (respectivement,  $M = 40,68, \acute{E}.-T. = 12,71$  et  $M = 34,44, \acute{E}.-T. = 11,97$ ) qu'à la première (respectivement,  $M = 35,84, \acute{E}.-T. = 13,67$  et  $M = 28,76, \acute{E}.-T. = 11,00$ ) ( $t(25) = -1,95, p = ,032$  et  $t(25) = -3,04, p = ,003$ ). De plus, ils utilisaient significativement plus de stratégies à la deuxième passation ( $M = 13,56, \acute{E}.-T. = 10,61$ ) qu'à la première ( $M = 8,64, \acute{E}.-T. = 7,77$ ) ( $t(25) = -3,63, p < ,001$ ). Il n'y avait pas de différence significative entre les deux temps de mesures quant au Nombre total de répétitions ( $t(25) = 0,82, p = ,419$ ) et au Pourcentage de répétitions ( $t(25) = 1,35, p = ,189$ ).

### **Équivalence**

Les scores au AppFLO et au FPT corrélaient fortement entre eux, tant pour le Nombre total de dessins ( $r = ,54, p < ,001$ ), le Nombre total de répétitions ( $r = ,58, p < ,001$ ), le Pourcentage de répétitions ( $r = ,68, p < ,001$ ), le Nombre de stratégies ( $r = ,62, p < ,001$ ) et le Nombre de dessins corrects ( $r = ,66, p < ,001$ ). Les scores de Pourcentage de dessins corrects aux deux versions du test corrélaient modérément entre eux ( $r = ,36, p = ,025$ ). Il n'y avait pas de différence significative entre les deux versions du test quant au Nombre total de dessins, au Nombre de dessins corrects et au Pourcentage de dessins corrects (Tableau 6). Cependant, les enfants effectuaient un nombre et un pourcentage significativement plus bas

de répétitions au FPT qu'au AppFLO et utilisaient significativement plus de stratégies au FPT qu'au AppFLO (Tableau 6).

**Tableau 6**

*Comparaison des performances au AppFLO et au FPT*

	Version du test		Valeur t
	AppFLO	FPT	
Nb total de dessins	34,87 ± 12,96	34,20 ± 10,40	0,32
Nb total de répétitions	5,73 ± 5,13	2,43 ± 2,50	4,29**
Nb de stratégies	7,97 ± 7,29	11,83 ± 10,62	-2,53*
Nb de dessins corrects	28,40 ± 10,37	28,20 ± 11,49	0,12
Pourcentage de répétitions	15,05 ± 12,08	7,81 ± 10,06	4,37**
Pourcentage de dessins corrects	82,91 ± 13,78	80,67 ± 17,40	0,69

Notes. \* =  $p < ,05$ , \*\* =  $p < ,01$ .

### Validité de construit

Les différents coefficients de corrélations obtenus entre les tests évaluant les fonctions exécutives et les scores principaux au AppFLO se trouvent dans le Tableau 7. Le Nombre de dessins corrects n'était pas corrélé au nombre de mots en fluidité verbale dans les conditions de lettres et sémantique, au temps d'initiation à la Tour de Londres, aux scores d'Empan spatial et au temps total à la condition 3 du Stroop. Le Nombre de stratégies au AppFLO était corrélé négativement au nombre de mouvements total à la Tour de Londres. De plus, le pourcentage de répétitions était corrélé négativement aux scores d'Empan spatial. Une corrélation quasi significative de taille modérée a aussi été obtenue entre le Nombre de dessins corrects au AppFLO, l'Empan spatial envers ( $p = ,051$ ) et le temps total condition 3 du Stroop ( $p = ,051$ ). À l'intérieur même du AppFLO, le Nombre de stratégies corrélait positivement et fortement avec le Nombre total de dessins, le Nombre de dessins corrects et le Nombre total de répétitions. Le Nombre total de dessins corrélaient également positivement et fortement avec le Nombre total de répétitions.

**Tableau 7**

*Coefficients de corrélation de Pearson entre les scores principaux au AppFLO et les scores en Fluidité verbale, à la Tour de Londres, au test d'Empan spatial et au Stroop*

	Nb total de dessins	Nb de dessins corrects	Nb de stratégies	Nb de répétitions	Pourc. de répétitions
	<i>r</i>				
Fluidité verbale – nb mots condition lettres	,05	,11	,10	-,07	-,13
Fluidité verbale – nb mots condition sémantique	,12	,19	,17	-,04	-,16
Fluidité verbale – nb mots condition alternance	,53**	,51**	,42*	,42*	,15
Fluidité verbale – nb alternance correcte	,51**	,55**	,50**	,33	,06
Tour de Londres – temps d'initiation	,03	-,02	,24	,17	,19
Tour de Londres – nb mouvements total	-,50**	-,51**	-,45*	-,33	-,09
Empan spatial endroit	,05	,21	,09	-,25	-,37*
Empan spatial envers	,20	,37 <sup>††</sup>	,21	-,16	-,38*
Stroop – temps total condition 3	-,32	-,37 <sup>††</sup>	-,25	-,11	-,01
Stroop – temps total condition 4	-,37*	-,49**	-,39*	-,02	,16
Stroop – erreurs total condition 3	-,20	-,40*	-,18	,22	,28
Stroop – erreurs total condition 4	,00	-,08	-,06	,17	,24
Nb total de dessins	1	,90**	,82**	,64**	,32
Nb de dessins corrects		1	,75**	,26	-,08
Nb de stratégies			1	,52**	,27
Nb de répétitions				1	,90**
Pourc. de répétitions					1

Notes. <sup>†</sup> = ,055 < p < ,065, <sup>††</sup> = ,050 < p < ,055, \* = p < ,050, \*\* = p < ,010.

Aucune des corrélations effectuées entre les différents scores au BRIEF (inhibition, alternance/flexibilité, contrôle émotionnel, initiation, mémoire de travail, planification, organisation et contrôle, IRC, IM et CEG) et les scores principaux au AppFLO ne s'est avérée significative ou près du seuil de signification.

### Discussion

Cette étude visait à vérifier la faisabilité d'une validation d'une version numérique d'un test de fluidité graphique inspiré du FPT, le AppFLO, chez des enfants de 6 à 12 ans, et d'observer les qualités psychométriques de l'outil. Ce test a été développé dans le but de répondre au besoin d'intégrer la technologie au sein de la neuropsychologie et de développer des outils d'évaluation neuropsychologique

adaptés à la réalité d'aujourd'hui. Le test a été administré à 30 enfants francophones âgés de 6 à 12 ans, ne présentant aucun trouble neurodéveloppemental, neurologique ou de santé mentale.

### **Faisabilité de l'étude**

Tout d'abord, les résultats indiquent que la totalité des participants a effectué l'ensemble des tâches prévues à la première rencontre d'évaluation. Cela laisse croire que la quantité de tâches prévues et le niveau de difficulté sont adéquats pour des enfants de 6 à 12 ans, suggérant une bonne faisabilité du projet. Le taux d'attrition de 6,67% est également acceptable (Nunan et al., 2018) et laisse entrevoir qu'il est accessible pour les participants de compléter les deux séances d'évaluation dans un délai de 14 à 28 jours. De plus, tous les enfants neurotypiques qui ont participé au projet ont complété l'entièreté de la tâche AppFLO, indiquant que le niveau de difficulté du test est adéquat et que celui-ci pourra être utilisé sans problème auprès d'autres enfants. D'ailleurs, les enfants rapportent un bon niveau d'appréciation de la tâche. En somme, la vérification de faisabilité de l'étude indique que le protocole de recherche ainsi que la tâche du AppFLO sont accessibles pour les participants.

### **Performances au AppFLO**

En ce qui concerne la relation entre les variables sociodémographiques ainsi que le QI et la performance au AppFLO, plusieurs de nos hypothèses ont été confirmées.

Nos résultats suggèrent que la performance des enfants au AppFLO s'améliore avec l'âge. En effet, comme attendu, les enfants plus âgés produisaient plus de dessins corrects, utilisaient plus de stratégies, et effectuaient un moindre pourcentage de répétitions relativement aux enfants plus jeunes. Certaines corrélations près du seuil de signification laissaient également entrevoir que les enfants plus âgés produisaient plus de dessins total, de dessins valides et un pourcentage plus élevé de dessins corrects que les enfants plus jeunes. Avec un échantillon plus grand, les résultats auraient possiblement atteint le seuil de la signification. Globalement, ces résultats concordent avec la littérature selon laquelle les FE se

développent graduellement au cours de l'enfance (Romine et Reynolds, 2005). En effet, bien que d'autres zones cérébrales seraient également impliquées lors de leur activation, les FE sont principalement associées au lobe préfrontal, qui serait l'une des dernières aires corticales à atteindre sa pleine maturité (Anderson, 2002; Lezak, 1995; Luciana, 2013; Lussier et al., 2017; Romine et Reynolds, 2005). Ainsi, les FE se développeraient graduellement et parallèlement au lobe frontal jusqu'à l'âge adulte (Romine et Reynolds, 2005). Plus précisément, les capacités d'inhibition connaîtraient une amélioration importante durant la période préscolaire (Best et al., 2009). Pour leur part, les habiletés en mémoire de travail et en flexibilité cognitive émergeraient durant la petite enfance, mais se développeraient de façon plus importante et linéaire durant la période scolaire (Best et al., 2009), suggérant que les améliorations observées dans la performance en FG pourraient s'expliquer par la croissance de ces fonctions. D'ailleurs, différentes études soulignent que la performance en FG s'améliorerait de manière graduelle durant l'enfance (Albert et al., 2009; Klenberg et al., 2001; Matute et al., 2004). En effet, les résultats sont similaires à ceux de la présente étude selon lesquels les enfants plus âgés produisaient plus de dessins corrects (Arffa, 2007; Klenberg et al., 2001; Stievano et Scalisi, 2016; Tucha et al., 2012) et utilisaient davantage de stratégies (Vik et Ruff, 1988). À notre connaissance, une seule étude s'est intéressée au lien entre l'âge et le pourcentage de répétitions pour la version papier du Five-Point Test chez les enfants. Les résultats indiquent l'absence de corrélation entre les deux variables (Stievano et Scalisi, 2016). Nos résultats indiquent, pour leur part, que le niveau scolaire, et non l'âge, était lié au pourcentage de répétitions. Lezak (1995) ainsi que Goebel et al. (2009) notent que l'utilisation de stratégies permettrait d'améliorer la performance au FPT, en augmentant le nombre de dessins produits et en diminuant le nombre de répétitions, puisqu'une série de dessins pourraient être produits à partir d'une idée. Selon nos résultats, les enfants plus âgés utiliseraient davantage de stratégies, ce qui pourrait expliquer du moins en partie que le pourcentage de répétitions diminuerait avec le niveau scolaire.

Au regard du sexe, nos résultats concernant les mesures standards de FG (Nombre total de dessins, de dessins corrects, de stratégies et de répétitions) sont en accord avec la littérature actuelle (Albert et al., 2009; Klenberg et al., 2001; Matias-Guiu et al., 2021; Regard et al., 1982; Van der Elst et al., 2011). En effet, ceux-ci suggèrent qu'il n'existe pas de différence entre la performance des filles et des garçons en fluidité graphique. Pour les autres mesures, telles que le nombre de dessins invalides ainsi que le nombre de répétitions adjacentes et non adjacentes, les résultats suggèrent également qu'il n'y aurait pas de différence entre les sexes. D'autres études sont toutefois requises afin de vérifier l'effet du sexe, la présente étude étant la première s'intéressant à ces mesures.

D'autre part, tel qu'observé par Arffa (2007), Goebel et al. (2009), Lee et al. (1997), Regard et al. (1982) et Vik and Ruff (1988), nos résultats suggèrent que le fonctionnement intellectuel serait lié à la performance en fluidité graphique. En effet, les enfants présentant une meilleure performance en raisonnement non verbal ont effectué davantage de dessins corrects, de stratégies et moins de dessins invalides au AppFLO. D'autre part, les enfants présentant une meilleure performance en raisonnement verbal ont produit davantage de dessins corrects et valides au AppFLO. Selon Duggan and Garcia-Barrera (2015), les fonctions exécutives et l'intelligence seraient deux construits fortement interreliés qui feraient appel à des processus cognitifs sous-jacents communs. Les deux fonctions seraient en partie sous-tendues par des aires cérébrales communes, soit les aires préfrontales et pariétales, ce qui expliquerait le lien observé (Duggan et Garcia-Barrera, 2015; Duncan et al., 2000; Funahashi, 2006). Ainsi, il n'est pas surprenant que différentes études obtiennent des résultats similaires à ceux de la présente étude selon lesquels les scores aux tests d'intelligence et de fonctionnement exécutif sont significativement corrélés entre eux (Ardila et al., 2000; Arffa, 2007; Brydges et al., 2012).

Contrairement aux résultats d'autres recherches (Klenberg et al., 2001; Van der Elst et al., 2011), aucune corrélation significative n'a été obtenue entre le niveau d'éducation des parents et les différents

scores en fluidité graphique. Cette divergence pourrait s'expliquer par le fait que notre échantillon était plus petit que dans les deux autres études. Il est aussi à noter que le niveau d'éducation des parents de notre échantillon était plus élevé que dans les autres études et n'était pas représentatif de la population québécoise. De plus, le niveau d'éducation des parents présentait peu de variabilité. En effet, la majorité des parents détenaient un diplôme universitaire. Il est possible qu'un niveau d'éducation élevé chez le parent n'entretienne pas de relation avec la performance de l'enfant en fluidité graphique et qu'avec un plus grand échantillon, dans lequel les différents niveaux d'éducation auraient été mieux représentés, une relation ait été obtenue.

### **Fidélité et validité**

En ce qui concerne les qualités psychométriques du test AppFLO, la présente étude laisse entrevoir une bonne fidélité et une bonne validité de l'outil.

*Fidélité test-retest.* D'abord, des degrés d'accord de modérés à élevés entre les deux temps de mesure suggèrent que la performance au test à deux moments rapprochés dans le temps est stable en ce qui a trait à la majorité des scores au AppFLO. Seuls les scores du nombre et du pourcentage de dessins invalides ne sont pas stables dans le temps. Ces variables présentaient peu de variabilité, puisque les erreurs de dessins invalides ne sont pas communes (Tucha et al., 2012), ce qui pourrait expliquer l'absence d'accord entre les deux temps de mesure. Malgré que les dessins invalides ne semblent pas être une mesure fidèle des fonctions exécutives dans la population tout venant, il demeure pertinent de les observer en clinique puisqu'elles pourraient être informatives qualitativement du fonctionnement cognitif de l'enfant ou de sa compréhension de la tâche chez différentes populations cliniques.

*Effet de pratique.* Un effet de pratique est remarqué au deuxième temps de passation du AppFLO. Les enfants effectuaient davantage de dessins total et corrects et utilisaient plus de stratégies lorsque le test leur a été administré la deuxième fois. Goebel et al. (2009) et Tucha et al. (2012) ont obtenu des

résultats similaires selon lesquels les participants produisaient plus de dessins corrects à la deuxième passation de la version papier du FPT. Dans leur étude de validation d'un autre test de fluidité graphique en version numérique, Woods et al. (2016) ont également observé cet effet de pratique. Notons que l'évaluation de la fidélité test-retest des tests mesurant les FE demeure un défi puisque l'un des éléments centraux de leur définition est que les FE sont activées lors de situations nouvelles (Lussier et al., 2017). Ainsi, la répétition de ce type de tâches pourrait faire en sorte que la performance soit davantage automatisée et ne dépende plus autant des habiletés exécutives, ce qui expliquerait l'effet de pratique observé (Hughes et Graham, 2002). Afin d'obtenir des résultats reflétant le plus précisément possible le fonctionnement exécutif de l'enfant, il n'est pas recommandé d'administrer le AppFLO à plusieurs reprises.

*Validité de construit – équivalence.* En ce qui concerne l'équivalence entre la version numérique et papier du test, les deux versions corrèlent fortement entre elles, et ce, pour les scores du nombre total de dessins, de dessins corrects, de stratégies, de répétitions et pour le pourcentage de répétitions. Pour le pourcentage de dessins corrects, les deux versions corrèlent modérément entre elles. Ces résultats suggèrent que les deux outils mesurent des construits similaires, soit le fonctionnement exécutif des enfants, et laisse entrevoir une bonne validité de construit du AppFLO. La version papier du FPT a un long historique et différentes études de validation soutiennent son utilité clinique (Fernandez et al., 2008; Goebel et al., 2009; Skalla, 2012; Stievano et Scalisi, 2016; Tucha et al., 2012). Ainsi, cette étude de validation du test AppFLO, qui démontre une forte validité de construit en termes d'équivalence entre les deux versions du test, permet de bénéficier de la réputation du test papier.

Cependant, les enfants effectuaient significativement plus de répétitions et moins de stratégies dans la version numérique du test. Rappelons que le test AppFLO a toujours été administré avant le FPT dans notre protocole d'évaluation. Il est donc possible que la différence quant au nombre et pourcentage de

répétitions ainsi qu'au nombre de stratégies s'explique en partie par un effet de pratique, tel qu'observé dans l'évaluation de la fidélité test-retest, les enfants ayant développé certaines stratégies et une automatisation de la tâche suite à la première passation du test en version numérique. Malgré cela, soulignons que dans la version numérique du test, les dessins préalablement effectués ne sont pas visibles, contrairement à la version papier du test. Ainsi, comme décrit par Woods et al. (2016), il se peut que l'augmentation du nombre de répétitions dans la version numérique du test soit aussi en partie expliquée par une demande mnésique plus élevée, en raison de l'impossibilité de se référer aux dessins produits préalablement. D'ailleurs, le pourcentage de répétitions moyen obtenu au FPT dans une étude de Stievano and Scalisi (2016) (8,79%) s'apparente au pourcentage de répétitions obtenu au FPT dans la présente étude (7,81%), bien qu'il n'y ait pas d'effet de pratique dans cette première étude, soutenant l'hypothèse que la différence observée dans la présente étude pourrait s'expliquer par la différence de modalité des tests (papier vs numérique), qui solliciteraient des fonctions différentes.

*Validité de construit – Convergence.* Par ailleurs, l'évaluation des corrélations entre le AppFLO et les tests mesurant des construits similaires et dont la validité est établie, soit la Fluidité verbale, la Tour de Londres, le test Empan spatial et le Stroop, laisse entrevoir une bonne validité convergente de l'outil. En effet, des corrélations de modérées à fortes ont été obtenues entre le AppFLO et ces autres outils mesurant les FE. D'abord, le nombre de dessins corrects au AppFLO corrèle fortement avec des mesures de flexibilité cognitive, de fluidité verbale, de planification et modérément avec des mesures d'inhibition et de mémoire de travail. Le nombre de dessins corrects semble ainsi se présenter comme une mesure globale des FE, qui impliquerait la coordination de différentes fonctions exécutives de base et de plus haut niveau recrutées durant la performance en fluidité graphique. Le nombre total de dessins corrèle pour sa part de modérément à fortement avec des mesures de flexibilité cognitive et de planification. Ensuite, le nombre de stratégies corrèle également de modérément à fortement avec des mesures de flexibilité

cognitive et de planification. D'autre part, le nombre total de répétitions corrèle modérément avec une seule mesure de flexibilité cognitive. Finalement, le pourcentage de répétitions corrèle modérément avec les mesures de mémoire de travail visuelle. Les mesures les plus intéressantes à considérer au test AppFLO semblent être le nombre de dessins corrects, qui serait une mesure globale des FE, le nombre de stratégies, qui évaluerait la flexibilité et la planification, ainsi que le pourcentage de répétitions, qui évaluerait la mémoire de travail visuelle. En somme, les fonctions cognitives évaluées par les différentes mesures du test AppFLO sont les mêmes que celles attendues, bien que les corrélations obtenues ne confirment pas l'ensemble des hypothèses émises (Stievano et Scalisi, 2016; Tucha et al., 2012). Plus précisément, le nombre de dessins corrects ne corrèle pas au nombre de mots dans les conditions de lettres et sémantique en fluidité verbale, au temps d'initiation à la Tour de Londres, aux scores d'Empan spatial et au temps total à la condition 3 du Stroop, comme attendu. Cependant, le nombre de dessins corrects au AppFLO corrèle avec d'autres mesures des mêmes tests, laissant entrevoir une bonne validité de construit malgré tout. Les différences observées pourraient en partie s'expliquer par le petit échantillon de la présente étude et la modalité d'administration du test. D'autres études seront nécessaires dans le but de clarifier les fonctions évaluées par les différentes mesures au AppFLO et vérifier si d'autres mesures, tel que les mesures de processus de complétion de la tâche ou les différents types de stratégies pourraient évaluer des fonctions exécutives plus spécifiques.

D'autre part, l'hypothèse concernant l'absence de relation entre les scores au AppFLO et au questionnaire BRIEF a été confirmée. La performance au test de fluidité graphique ne corrèle pas avec les échelles de l'évaluation comportementale. À notre connaissance, aucune étude ne s'est intéressée au lien entre la performance aux tests en fluidité graphique et les scores au BRIEF. Cependant, Anderson et al. (2002), Mahone et al. (2002), Mangeot et al. (2002) ainsi que Niendam et al. (2007) ont évalué la corrélation entre différents tests de fluidité verbale et le BRIEF. La majorité des études n'observent pas

de corrélation entre les outils. Lorsqu'un lien est observé, il est de petit effet (Anderson et al., 2002). Selon Toplak et al. (2013), l'absence de corrélation entre les évaluations comportementales de FE et les tests basés sur la performance s'expliquerait par le fait qu'ils mesurent des aspects différents des FE : les questionnaires évalueraient davantage la poursuite et l'atteinte des objectifs individuels alors que les tests mesureraient l'efficacité et l'intégrité des processus cognitifs. Ainsi, les deux types d'outils sont pertinents, mais fournissent de l'information différente, suggérant qu'il peut être utile en clinique d'administrer les deux afin d'obtenir un portrait plus global des FE. Par ailleurs, l'absence de corrélation entre le BRIEF et le AppFLO pourrait être considérée comme un indice supplémentaire de la validité de construit de notre test.

Finalement, les corrélations entre les scores au AppFLO sont concordantes avec les hypothèses émises et sont un autre indicateur de la validité de construit de l'outil. Comme attendu, l'utilisation d'un plus grand nombre de stratégies favorisait la productivité et la précision en augmentant le nombre de dessins total et corrects. De plus, comme attendu, un plus grand nombre total de dessins était lié à un plus grand nombre total de répétitions. Toutefois, les enfants qui effectuaient plus de dessins total ne présentaient pas un taux plus élevé de répétitions, suggérant que la qualité de leur production globale n'était pas affectée. Concernant la relation entre le nombre de stratégies et le nombre total de répétitions, nos résultats diffèrent de ceux obtenus par Stievano and Scalisi (2016) et Goebel et al. (2009). Dans notre étude, l'utilisation d'un plus grand nombre de stratégies était liée à un nombre total de répétitions plus élevé. Cette relation peut possiblement s'expliquer par le fait qu'une utilisation accrue de stratégies permet d'effectuer plus de dessins total, ce qui engendre un plus grand nombre de répétitions. Malgré tout, la qualité de la production globale ne semble pas affectée par le nombre de stratégies puisque nous n'observons pas de corrélation entre cette mesure et le pourcentage de répétitions.

En somme, cette étude pilote nous indique une bonne faisabilité du projet de validation du AppFLO, en termes d'acceptabilité de l'outil par les enfants et du taux d'attrition des participants. De plus, la prévalidation suggère que le test présente de bonnes qualités psychométriques. En effet, les corrélations modérées à élevées entre les deux temps de passation de l'outil laissent voir une bonne fidélité test-retest. Un effet de pratique est cependant observé, les enfants produisant davantage de dessins total et corrects et utilisant plus de stratégies à la deuxième passation. D'autre part, les corrélations modérées à élevées entre le FPT et AppFLO en plus des corrélations modérées à élevées entre les tests de FE et AppFLO indiquent une validité de construit adéquate. Par ailleurs, les résultats suggèrent que les mesures qui semblent les plus intéressantes à considérer au test AppFLO seraient le nombre de dessins corrects, qui serait une mesure globale des FE, le nombre de stratégies, qui évaluerait la flexibilité et la planification, ainsi que le pourcentage de répétitions, qui évaluerait la mémoire de travail visuelle. Finalement, les corrélations entre les scores à l'intérieur du test sont concordantes avec les hypothèses émises et sont un autre indicateur de la validité de l'outil.

### **Limites**

En raison de la nature préliminaire de cette étude, certaines limitations méthodologiques doivent être prises en considération dans les interprétations. Tout d'abord, la taille restreinte de l'échantillon a pu affecter la puissance statistique et limiter les interprétations que nous pouvons tirer des résultats. En raison du petit nombre de participants, il n'a d'ailleurs pas été possible d'effectuer des analyses comparant la performance des enfants en fonction de leur âge ou de développer des normes préliminaires pour l'outil en fonction des groupes d'âge. D'autre part, la représentativité de l'échantillonnage est une limite de la présente étude. L'échantillon est composé de davantage de filles que de garçons. De plus, les différents niveaux scolaires ne sont pas représentés de manière égale, un seul jeune de 6<sup>e</sup> année étant présent. Il n'y a aussi qu'un enfant âgé de 6 ans et un seul âgé de 12 ans. Par ailleurs, les enfants proviennent tous de

Montréal ou de Québec, les parents possèdent majoritairement un diplôme universitaire et ont majoritairement un revenu annuel familial brut au-dessus de 100 000\$, ce qui n'est pas représentatif de l'ensemble de la population québécoise et limite la généralisation des résultats. Finalement, le protocole selon lequel le AppFLO était toujours administré avant le FPT n'était pas optimal. En effet, il n'est pas possible de départager si les différences entre les deux versions du test sont explicables par la modalité de passation qui diffère ou simplement par un effet de pratique. Il aurait donc été souhaitable, pour évaluer la validité de construit selon la méthode d'équivalence, de contrebalancer la passation des deux outils, mais au prix d'un plus grand échantillon pour satisfaire aux autres mesures de validation.

### **Conclusion**

En somme, il s'agit, à notre connaissance, de la première étude visant à valider une version numérique d'un test de fluidité graphique sur tablette chez des enfants d'âge scolaire. Bien que le petit échantillon limite certaines des conclusions, il demeure que les résultats préliminaires quant à la validité et la fidélité de l'outil sont encourageants et portent à croire qu'une validation plus large du test et la création de normes sont pertinentes pour le futur. Ainsi, les prochaines étapes dans le développement de l'outil seront d'effectuer d'autres études avec des échantillons plus grands et plus représentatifs de la population québécoise dans le but de mieux documenter les qualités psychométriques et développer des normes qui pourront être utilisées à des fins d'évaluation. Des études cliniques pourront également être effectuées auprès de diverses populations pédiatriques afin d'en évaluer l'utilité pour caractériser le fonctionnement exécutif et détecter les dysfonctions exécutives, par exemple dans le TDA/H, le trouble du spectre de l'autisme, le trouble développemental du langage et les traumatismes craniocérébraux. L'outil pourra ensuite être validé et normé auprès d'adultes et d'ainées pour en élargir l'utilisation dans différentes populations cliniques incluant démences, aphasies, amnésies et maladie de Parkinson. Finalement, cette étude contribue à une meilleure connaissance du lien entre la performance en fluidité

graphique et les différents facteurs sociodémographiques, en plus de contribuer à la modernisation des pratiques en neuropsychologie clinique et au développement d'outils d'évaluation numériques.

## Références

- Albert, D., Opwis, K., & Regard, M. (2009). Effect of drawing hand and age on figural fluency: a graphomotor study with the five-point test in children. *Child Neuropsychol*, 16(1), 32-41. <https://doi.org/https://doi.org/10.1080/09297040903049061>
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders : DSM-5* (Vol. 5).
- Anderson, P. (2002). Assessment and Development of Executive Function (EF) During Childhood. *Child Neuropsychology*, 8(2), 71-82. <https://doi.org/10.1076/chin.8.2.71.8724>
- Anderson, V. A., Anderson, P., Northam, E., Jacobs, R., & Mikiewicz, O. (2002). Relationships Between Cognitive and Behavioral Measures of Executive Function in Children With Brain Disease. *Child Neuropsychology*, 8(4), 231-240. <https://doi.org/10.1076/chin.8.4.231.13509>
- Ardila, A., Pineda, D., & Rosselli, M. (2000). Correlation between intelligence test scores and executive function measures. *Arch Clin Neuropsychol*, 15(1), 31-36.
- Arffa, S. (2007). The relationship of intelligence to executive function and non-executive function measures in a sample of average, above average, and gifted youth. *Arch Clin Neuropsychol*, 22(8), 969-978. <https://doi.org/10.1016/j.acn.2007.08.001>
- Baldo, J. V., Shimamura, A. P., Delis, D. C., Kramer, J., & Kaplan, E. (2001). Verbal and design fluency in patients with frontal lobe lesions. *Journal of the International Neuropsychological Society*, 7(5), 586-596. <https://doi.org/10.1017/S1355617701755063>
- Best, J. R., Miller, P. H., & Jones, L. L. (2009). Executive functions after age 5: Changes and correlates. *Developmental Review*, 29(3), 180-200. <https://doi.org/https://doi.org/10.1016/j.dr.2009.05.002>
- Bilder, R. M., & Reise, S. P. (2019). Neuropsychological tests of the future: How do we get there from here? *The Clinical Neuropsychologist*, 33(2), 220-245. <https://doi.org/10.1080/13854046.2018.1521993>
- Browne, R. H. (1995). On the use of a pilot sample for sample size determination. *Stat Med*, 14(17), 1933-1940. <https://doi.org/10.1002/sim.4780141709>
- Brydges, C. R., Reid, C. L., Fox, A. M., & Anderson, M. (2012). A unitary executive function predicts intelligence in children. *Intelligence*, 40(5), 458-469. <https://doi.org/https://doi.org/10.1016/j.intell.2012.05.006>
- Cattelani, R., Dal Sasso, F., Corsini, D., & Posteraro, L. (2011). The Modified Five-Point Test: normative data for a sample of Italian healthy adults aged 16–60. *Neurological Sciences*, 32(4), 595-601. <https://doi.org/10.1007/s10072-011-0489-4>
- Cortés Pascual, A., Moyano Muñoz, N., & Quílez Robres, A. (2019). The Relationship Between Executive Functions and Academic Performance in Primary Education: Review and Meta-Analysis

[Systematic Review]. *Frontiers in Psychology*, 10(1582).  
<https://doi.org/10.3389/fpsyg.2019.01582>

- Culbertson, W. C., & Zillmer, E. A. (1998). The Tower of LondonDX: A Standardized Approach to Assessing Executive Functioning in Children. *Archives of Clinical Neuropsychology*, 13(3), 285-301. [https://doi.org/10.1016/S0887-6177\(97\)00033-4](https://doi.org/10.1016/S0887-6177(97)00033-4)
- Culbertson, W. C., Zillmer, E. A., Panchal, P., & Giordano, H. (1997). Preliminary childhood norms for the TOLDX. *Archives of Clinical Neuropsychology*, 12(4), 304-305. <https://doi.org/10.1093/arclin/12.4.304>
- Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). Delis-Kaplan executive function system.
- Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, 64(1), 135-168. <https://doi.org/10.1146/annurev-psych-113011-143750>
- Duggan, E. C., & Garcia-Barrera, M. A. (2015). Executive functioning and intelligence. *Handbook of intelligence: Evolutionary theory, historical perspective, and current concepts*, 435-458.
- Duncan, J., Seitz, R., Kolodny, J., Bor, D., Herzog, H., Ahmed, A., Newell, F., & Emslie, H. (2000). A neural basis for general intelligence. *Science (New York, N.Y.)*, 289, 457-460. [https://doi.org/10.1016/S0002-9394\(00\)00752-2](https://doi.org/10.1016/S0002-9394(00)00752-2)
- Fernandez, A., Moroni, M., Carranza, J., Fabbro, N., & Lebowitz, B. (2008). Reliability of the Five-Point Test. *The Clinical Neuropsychologist*, 23, 501-509. <https://doi.org/10.1080/13854040802279675>
- Funahashi, S. (2006). Prefrontal cortex and working memory processes. *Neuroscience*, 139(1), 251-261. <https://doi.org/10.1016/j.neuroscience.2005.07.003>
- Gauthier, B., Parent, V., & Lageix, P. (2016). Exploring the dynamics of design fluency in children with and without ADHD using artificial neural networks. *Child Neuropsychology*, 22(2), 238-246. <https://doi.org/10.1080/09297049.2014.988606>
- Ghanavati, E., Salehinejad, M. A., Nejati, V., & Nitsche, M. A. (2019). Differential role of prefrontal, temporal and parietal cortices in verbal and figural fluency: Implications for the supramodal contribution of executive functions. *Scientific Reports*, 9(1), 3700. <https://doi.org/10.1038/s41598-019-40273-7>
- Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *Behavior rating inventory of executive function: BRIEF*. Psychological Assessment Resources Odessa, FL.
- Goebel, S., Fischer, R., Ferstl, R., & Mehdorn, H. M. (2009). Normative data and psychometric properties for qualitative and quantitative scoring criteria of the Five-point Test. *Clin Neuropsychol*, 23(4), 675-690. <https://doi.org/10.1080/13854040802389185>
- Holmes, C. J., Kim-Spoon, J., & Deater-Deckard, K. (2016). Linking Executive Function and Peer Problems from Early Childhood Through Middle Adolescence. *Journal of abnormal child psychology*, 44(1), 31-42. <https://doi.org/10.1007/s10802-015-0044-5>

- Hughes, C., & Graham, A. (2002). Measuring executive functions in childhood: Problems and solutions? *Child and adolescent mental health*, 7(3), 131-142.
- Hurks, P. P. M. (2013). Administering Design Fluency Tests in School-aged Children: Analyses of Design Productivity over Time, Clustering, and Switching. *The Clinical Neuropsychologist*, 27(7), 1131-1149. <https://doi.org/10.1080/13854046.2013.821170>
- Julious, S. A. (2005). Sample size of 12 per group rule of thumb for a pilot study. *Pharmaceutical Statistics*, 4(4), 287-291. <https://doi.org/https://doi.org/10.1002/pst.185>
- Klenberg, L., Korkman, M., & Lahti-Nuutila, P. (2001). Differential development of attention and executive functions in 3- to 12-year-old Finnish children. *Dev Neuropsychol*, 20(1), 407-428. [https://doi.org/10.1207/s15326942dn2001\\_6](https://doi.org/10.1207/s15326942dn2001_6)
- Kwak, S. G., & Kim, J. H. (2017). Central limit theorem: the cornerstone of modern statistics. *kja*, 70(2), 144-156. <https://doi.org/10.4097/kjae.2017.70.2.144>
- Lee, G. P., Strauss, E., Loring, D. W., McCloskey, L., Haworth, J. M., & Lehman, R. A. W. (1997). Sensitivity of figural fluency on the five-point test to focal neurological dysfunction. *The Clinical Neuropsychologist*, 11(1), 59-68. <https://doi.org/10.1080/13854049708407030>
- Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, 21(1), 59-80. <https://doi.org/10.1348/026151003321164627>
- Levin, H. S., Culhane, K. A., Hartmann, J., Evankovich, K., Mattson, A. J., Harward, H., Ringholz, G., Ewing-Cobbs, L., & Fletcher, J. M. (1991). Developmental changes in performance on tests of purported frontal lobe functioning. *Developmental Neuropsychology*, 7(3), 377-395. <https://doi.org/10.1080/87565649109540499>
- Lezak, M. D. (1995). *Neuropsychological assessment, 3rd ed.* Oxford University Press.
- Luciana, M. (2013). Adolescent brain development in normality and psychopathology. *Development and Psychopathology*, 25(4pt2), 1325-1345. <https://doi.org/10.1017/S0954579413000643>
- Lussier, F., Chevrier, E., & Gascon, L. (2017). *Neuropsychologie de l'enfant et de l'adolescent: Troubles développementaux et de l'apprentissage.* Dunod. <https://books.google.ca/books?id=vmqstAEACAAJ>
- Mahone, E. M., Cirino, P. T., Cutting, L. E., Cerrone, P. M., Hagelthorn, K. M., Hiemenz, J. R., Singer, H. S., & Denckla, M. B. (2002). Validity of the behavior rating inventory of executive function in children with ADHD and/or Tourette syndrome. *Arch Clin Neuropsychol*, 17(7), 643-662.
- Mangeot, S., Armstrong, K., Colvin, A. N., Yeates, K. O., & Taylor, H. G. (2002). Long-term executive function deficits in children with traumatic brain injuries: assessment using the Behavior Rating Inventory of Executive Function (BRIEF). *Child Neuropsychol*, 8(4), 271-284. <https://doi.org/10.1076/chin.8.4.271.13503>

- Matias-Guiu, J. A., Pytel, V., Delgado-Álvarez, A., Delgado-Alonso, C., Cortés-Martínez, A., Fernández-Oliveira, A., & Matías-Guiu, J. (2021). The Five-Point Test: Normative data for middle-aged and elderly Spaniards. *Appl Neuropsychol Adult*, 1-9. <https://doi.org/10.1080/23279095.2021.1873137>
- Matute, E., Rosselli, M., Ardila, A., & Morales, G. (2004). Verbal and Nonverbal Fluency in Spanish-Speaking Children. *Developmental Neuropsychology*, 26(2), 647-660. [https://doi.org/10.1207/s15326942dn2602\\_7](https://doi.org/10.1207/s15326942dn2602_7)
- Miller, J. B., & Barr, W. B. (2017). The Technology Crisis in Neuropsychology. *Archives of Clinical Neuropsychology*, 32(5), 541-554. <https://doi.org/10.1093/arclin/acx050>
- 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. *Cognitive Psychology*, 41(1), 49-100. <https://doi.org/https://doi.org/10.1006/cogp.1999.0734>
- Nejati, V., Salehinejad, M. A., & Nitsche, M. A. (2018). Interaction of the Left Dorsolateral Prefrontal Cortex (l-DLPFC) and Right Orbitofrontal Cortex (OFC) in Hot and Cold Executive Functions: Evidence from Transcranial Direct Current Stimulation (tDCS). *Neuroscience*, 369, 109-123. <https://doi.org/https://doi.org/10.1016/j.neuroscience.2017.10.042>
- Niendam, T. A., Horwitz, J., Bearden, C. E., & Cannon, T. D. (2007). Ecological assessment of executive dysfunction in the psychosis prodrome: A pilot study. *Schizophrenia Research*, 93(1), 350-354. <https://doi.org/https://doi.org/10.1016/j.schres.2007.03.009>
- Nunan, D., Aronson, J., & Bankhead, C. (2018). Catalogue of bias: attrition bias. *BMJ Evid Based Med*, 23(1), 21-22. <https://doi.org/10.1136/ebmed-2017-110883>
- Rabin, L. A., Paolillo, E., & Barr, W. B. (2016). Stability in Test-Usage Practices of Clinical Neuropsychologists in the United States and Canada Over a 10-Year Period: A Follow-Up Survey of INS and NAN Members. *Arch Clin Neuropsychol*, 31(3), 206-230. <https://doi.org/10.1093/arclin/acw007>
- Rabin, L. A., Spadaccini, A. T., Brodale, D. L., Grant, K. S., Elbulok-Charcape, M. M., & Barr, W. B. (2014). Utilization rates of computerized tests and test batteries among clinical neuropsychologists in the United States and Canada. *Professional Psychology: Research and Practice*, 45(5), 368-377. <https://doi.org/10.1037/a0037987>
- Regard, M., Strauss, E., & Knapp, P. (1982). Children's Production on Verbal and Non-Verbal Fluency Tasks. *Perceptual and Motor Skills*, 55(3), 839-844. <https://doi.org/10.2466/pms.1982.55.3.839>
- Reiter, A., Tucha, O., & Lange, K. W. (2005). Executive functions in children with dyslexia [<https://doi.org/10.1002/dys.289>]. *Dyslexia*, 11(2), 116-131. <https://doi.org/https://doi.org/10.1002/dys.289>
- Riggs, N. R., Blair, C. B., & Greenberg, M. T. (2004). Concurrent and 2-Year Longitudinal Relations Between Executive Function and the Behavior of 1st and 2nd Grade Children. *Child Neuropsychology*, 9(4), 267-276. <https://doi.org/10.1076/chin.9.4.267.23513>

- Robinson, T., & Tripp, G. (2013). Neuropsychological functioning in children with ADHD: Symptom persistence is linked to poorer performance on measures of executive and nonexecutive function. *Japanese Psychological Research*, 55. <https://doi.org/10.1111/jpr.12005>
- Romine, C. B., & Reynolds, C. R. (2005). A Model of the Development of Frontal Lobe Functioning: Findings From a Meta-Analysis. *Applied Neuropsychology*, 12(4), 190-201. [https://doi.org/10.1207/s15324826an1204\\_2](https://doi.org/10.1207/s15324826an1204_2)
- Ruff, R. M. (1996). *RFFT: Ruff figural fluency test: Professional manual*. PAR.
- Schultheis, M. T., & Doiron, M. (2017). Technologies for Functionally Relevant Neuropsychological Assessment. In R. L. Kane & T. D. Parsons (Eds.), *The Role of Technology in Clinical Neuropsychology*. (pp. 47-60). Oxford University Press. <https://doi.org/oso/9780190234737.003.0003>
- Skalla, J. (2012). Using the Quantified Process Approach in Examination of the Five Point Test. *ETD Archive*, 609. <https://doi.org/https://engagedscholarship.csuohio.edu/etdarchive/609/>
- Stievano, P., & Scalisi, T. G. (2016). Unique designs, errors and strategies in the Five-Point Test: The contribution of age, phonemic fluency and visuospatial abilities in Italian children aged 6-11 years. *Child Neuropsychol*, 22(2), 197-219. <https://doi.org/10.1080/09297049.2014.988607>
- Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary, 3rd ed.* Oxford University Press.
- Suchy, Y., Kraybill, M. L., & Gidley Larson, J. C. (2010). Understanding design fluency: Motor and executive contributions. *Journal of the International Neuropsychological Society*, 16(1), 26-37. <https://doi.org/10.1017/S1355617709990804>
- 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? *Journal of Child Psychology and Psychiatry*, 54(2), 131-143. <https://doi.org/https://doi.org/10.1111/jcpp.12001>
- Tucha, L., Aschenbrenner, S., Koerts, J., & Lange, K. W. (2012). The Five-Point Test: Reliability, Validity and Normative Data for Children and Adults. *PLOS ONE*, 7(9), e46080. <https://doi.org/10.1371/journal.pone.0046080>
- Van der Elst, W., Hurks, P., Wassenberg, R., Meijs, C., & Jolles, J. (2011). Animal Verbal Fluency and Design Fluency in school-aged children: Effects of age, sex, and mean level of parental education, and regression-based normative data. *Journal of Clinical and Experimental Neuropsychology*, 33(9), 1005-1015. <https://doi.org/10.1080/13803395.2011.589509>
- Vélez-van-Meerbeke, A., Zamora, I. P., Guzmán, G., Figueroa, B., López Cabra, C. A., & Talero-Gutiérrez, C. (2013). Evaluación de la función ejecutiva en una población escolar con síntomas de déficit de atención e hiperactividad. *Neurología*, 28(6), 348-355. <https://doi.org/https://doi.org/10.1016/j.nrl.2012.06.011>

- Vik, P., & Ruff, R. R. (1988). Children's figural fluency performance: Development of strategy use. *Developmental Neuropsychology*, 4(1), 63-74. <https://doi.org/10.1080/87565648809540391>
- Wang, Y., & Liu, Y. (2021). *The development of internalizing and externalizing problems in primary school: Contributions of executive function and social competence* [doi:10.1111/cdev.13462]. Wiley-Blackwell Publishing Ltd.
- Wang, Y., & Zhou, X. (2019). Longitudinal relations between executive function and internalizing problems in grade school: The role of peer difficulty and academic performance [doi:10.1037/dev0000790]. 55, 2147-2158. <https://doi.org/10.1037/dev0000790>
- Wechsler, D. (2015). *Échelle d'intelligence de Wechsler pour enfants – Cinquième édition : Version francophone du Canada (WISC-V CDN-F)*. Pearson Canada Assessment, Inc.
- Wechsler, D., Kaplan, E., Fein, D., Kramer, J., Morris, R., Delis, D., & Maerlander, A. (2004). *WISC-IV Integrated*. Wechsler Intelligence Scale for Children.
- Woods, D. L., Wyma, J. M., Herron, T. J., & Yund, E. W. (2016). A Computerized Test of Design Fluency. *PloS one*, 11(5), e0153952-e0153952. <https://doi.org/10.1371/journal.pone.0153952>