

Test of Everyday Attention, Second Edition – Review (janvier 2019)

Par Julie Duval, Sébastien Monette et Josée Douaire

La deuxième édition du Test of Everyday Attention (TEA-Ch2; Manly, Anderson, Crawford, George et Robertson, 2016) est vendue au Québec par Pearson depuis le début de l'année 2018. Cette nouvelle version s'adresse aux enfants de 5 à 15 ans. La première partie de la batterie est administrée à l'aide de protocoles papier et la seconde partie est administrée à l'ordinateur. La batterie anglophone n'a pas été traduite, mais les stimuli sont entièrement non verbaux, ce qui permet une adaptation aisée de l'administration en français par la simple traduction des consignes.

La batterie est scindée en deux versions selon l'âge de l'enfant. La première version s'adresse aux jeunes de 5 à 7 ans (junior - J) et la seconde aux jeunes de 8 à 15 ans (adolescent - A).

Description des sous-tests

La batterie contient 7 sous-tests dans la version junior (J) et 9 dans la version adolescent (A). Le temps d'administration estimé est de 35-40 minutes pour la version junior et 40-55 minutes pour la version adolescent. La batterie donne un seul score par sous-test et les résultats sont offerts en cotes standards.

1- Balloon Hunt (J) / Hector Cancelation (A)

Il s'agit d'une épreuve classique d'annulation, ressemblant au barrage de cloches, mais avec une limite de temps. L'enfant doit faire un trait sur le plus grand nombre de cibles possible dans le temps alloué (J=15 sec, A=10 sec.). On répète l'exercice sur plusieurs pages (J=4 essais, A=6 essais).

Score : nombre moyen de cibles identifiées aux essais combinés.

2- Balloons 5 (J) / Hector B Cancelation (A)

Un essai supplémentaire d'annulation est proposé, mais sans limite de temps cette fois. On demande à l'enfant de trouver toutes les cibles et on lui indique qu'il peut prendre tout le temps nécessaire (donc on ne demande pas une exécution rapide).

Score : ratio du temps par cible (nombre de cibles identifiées divisé par le temps d'exécution).

3 - Hide & Seek Visual (J) / Hecuba Visual Search (A)

Il s'agit encore d'une épreuve de recherche visuelle, mais cette fois la réponse est donnée verbalement par l'enfant. Une série d'encadrés est présentée sur une feuille et l'enfant doit dire si oui ou non la cible recherchée est présente dans chacun des encadrés. L'enfant doit compléter le plus d'items possible dans le temps alloué (J=60 sec., A=30 sec.) sur deux essais.

Score: nombre moyen de réponses correctes aux deux essais.

4- Troy Dual Task (A)

Épreuve d'attention partagée. L'enfant doit effectuer à nouveau une épreuve d'annulation en même temps qu'il doit compter le nombre de stimuli sonores entendus. Quatre essais sont proposés qui consistent chacun en une feuille de recherche visuelle pour une durée de 10 secondes durant lesquelles les stimuli sonores sont entendus. Les stimuli sonores sont aisément comptés (jamais plus de 7), si bien que les auteurs mentionnent qu'il est peu typique de faire des erreurs de décompte à tout âge. D'ailleurs, les erreurs de dénombrement ne sont pas recensées par le logiciel ou comptabilisées dans le résultat comme tel.

Score: nombre moyen de cibles visuelles correctement identifiées aux 4 essais combinés.

5- Barking (J) / Vigil (A)

Épreuve qui ressemble aux 'coups de fusil' de la version originale. L'enfant doit tenir le décompte des stimuli entendus. Un essai sur deux présente des stimuli à intervalles réguliers et l'autre à intervalles irréguliers. Contrairement au test de la version originale, on permet à l'enfant de compter à voix haute. Le décompte ne dépasse jamais 6 dans la version junior, alors qu'il monte à 14 dans la version adolescent.

Score : nombre de réponses correctes sur 10 essais.

6- Hide & Seek Auditory (J) / Cerberus (A)

Tâche de détection en modalité auditive. L'enfant doit réagir à un son-cible entendu parmi d'autres distracteurs sonores. Le son-cible survient soit une fois par essai ou pas du tout (J=10 essais, A=15 essais). L'enfant doit appuyer sur la barre d'espace le plus rapidement possible lorsqu'il entend le son-cible et à la fin de l'essai, il indique si oui ou non la cible a été entendue durant l'essai.

Score : temps de réaction moyen aux cibles auditives lorsque la cible est présente (appuyer sur la barre d'espace) divisé par le nombre total de réponses correctes. Une réponse est considérée correcte lorsque: l'enfant dit qu'il a entendu la cible à la fin d'un item (oui), la cible était présente et il a appuyé sur la barre d'espace durant l'item.

7- Simple RT (J et A)

On demande à l'enfant d'appuyer sur la barre d'espace à l'apparition d'un cercle bleu à l'écran (il peut apparaître au centre, à droite ou à gauche à une vitesse aléatoire). Il n'y a aucun distracteur et la cible reste visible tant que la réponse n'a pas été donnée (dans la limite de temps offerte). Si l'enfant fait une erreur (commission ou omission), un point d'interrogation apparaît à l'écran et l'examineur doit recadrer l'enfant et cliquer sur une

icône à l'écran afin de reprendre l'administration du sous-test (le temps de réaction ne sera pas comptabilisé pour cette cible).

Score : temps de réaction moyen aux bonnes réponses.

8- SART (J et A)

Épreuve similaire au CPT. Des formes géométriques simples apparaissent une à la fois au centre de l'écran (cercle, carré, triangle). L'enfant doit appuyer sur la barre d'espace après chacune d'elle, à l'exception du triangle. À la différence du CPT, il n'y a pas de composante de vitesse puisque l'enfant est invité à appuyer sur la barre d'espace suite à l'apparition d'une flèche qui descend vers le bas, présentée immédiatement après la forme.

Score : nombre total de réponses faites aux items 'no-go' (erreur de commission).

9- RBBS (A)

Épreuve d'alternance. L'enfant doit appairer des stimuli en fonction de 2 concepts, soit la couleur (rouge ou bleu) et la partie du corps (se tient dans la main ou se porte aux pieds). Ainsi, un soulier rouge sera associé à la couleur rouge dans la première condition, alors qu'il sera associé au pied dans la seconde. Après une pratique pour chaque condition, on propose deux essais d'alternance en passant d'une condition à l'autre à plusieurs reprises.

Score : le temps de réaction moyen aux réponses correctes pour les deux essais d'alternance.

Cotation

La correction se fait par le logiciel (aucun tableau de norme disponible dans le manuel). Le score brut des sous-tests est converti en score pondéré ($M = 10$, $\text{É.-T.} = 3$). Trois scores composites sont offerts sous forme de scores standards (moyenne 100, écart-type de 15) : un score composite d'attention sélective (*Selective Attention Index*), un score composite d'attention soutenue (*Sustained Attention Index*) et un score global (*Everyday Attention Index*).

Normes et psychométrie

Les normes ont été recueillies au Royaume-Uni. L'échantillon se compose de 394 enfants pour la version junior et 621 pour la version adolescent. L'échantillon est représentatif de la population du Royaume-Uni à l'égard du sexe et de l'ethnie de l'enfant, du niveau d'éducation familial et de la région de provenance. La cotation peut se faire sur la base de l'échantillon total ou en prenant en considération le niveau d'éducation familiale. Par contre, ce dernier aspect est difficile à arrimer avec notre système d'éducation québécois.

Fiabilité

Consistance interne. Les indices de consistance interne sont variables d'un sous-test à l'autre, passant d'excellents ($>.90$) à modérés ($>.50$). Pour l'échantillon normatif, la

majorité des sous-tests de la version J (5/7) et de la version A (7/9) montrent une consistance interne acceptable ou mieux ($>.70$). Les indices de consistance interne sont généralement un peu plus élevés dans l'échantillon clinique.

Fidélité test-retest. La fidélité test-retest fluctue également, variant de faible (.38) à bon (.84). Pour la version J seulement 2/7 sous-tests montrent une corrélation test-retest acceptable ou mieux ($r >.70$) et 5/9 sous-tests pour la version A. La fidélité test-retest n'est donc pas équivalente pour tous les sous-tests de cette batterie, étant plus faible pour la version junior. Mais tel qu'attendu, les indices de fiabilité sont plus élevés au niveau des scores composites.

Validité

Validité factorielle. Au niveau de la structure interne, les analyses factorielles confirmatoires supportent un modèle à 2 facteurs corrélés, soit l'attention sélective et l'attention soutenue. Les poids factoriels de certains sous-tests sont plutôt faibles ($<.40$) pour 4/7 sous-tests de la version J et 4/9 pour la version A, ce qui suggère que le facteur général d'attention auquel le sous-test est associé (sélective ou soutenue) explique peu de variance pour certains sous-tests.

Validité convergente. Les études corrélationnelles avec d'autres tests ont porté sur les résultats à un questionnaire de comportement (observations parentales) et au WISC-IV.

- *Strengths and Difficulties Questionnaire*. Les auteurs rapportent les corrélations entre les sous-tests et échelles du TEA-Ch2 et les échelles du SDQ (problèmes émotionnels, sociaux, de conduite, de TDAH et comportements prosociaux) dans cinq tableaux (version J+SDQ parent, version J+SDQ enseignant, version A+SDQ parent, version A+SDQ enseignant, version A+SDQ auto-rapporté). Les corrélations du score global au TEA-Ch2 et de l'échelle TDAH du SDQ pour les cinq groupes précédemment nommés sont : -.24, -.15, -.08, .12 et -.03, suggérant une faible validité convergente avec les symptômes du TDAH.
- WISC-IV^{UK} version courte (4 sous-tests : Vocabulaire, Similitudes, Blocs et Matrices). Les corrélations entre les sous-tests du TEA-Ch2 et les sous-tests du WISC-IV sélectionnées sont de faibles à modérées ($r = -.07$ à $.38$) ce qui semble confirmer que l'influence du niveau intellectuel reste limitée, mais sans être inexistante puisque certains corrélent plus fortement, notamment avec le sous-test matrices et les indices globaux (IAG, IRF, ICV). Concernant le score global du TEA-Ch2, il corréle modérément avec les sous-tests du WISC-IV ($r = .20$ à $.45$).

Études cliniques/validité discriminante

Les auteurs du test nous offrent des résultats à trois études cliniques.

Dans la première étude, le groupe clinique est composé de 139 enfants référés pour un problème d'attention (on ne sait donc pas s'ils ont réellement un TDAH ou s'ils ont une

autre problématique), lesquels sont appariés à un groupe contrôle dérivé de l'échantillon normatif et apparié pour le genre, l'âge et le niveau d'éducation familiale. L'étude ne donne pas des résultats pour l'ensemble des sous-tests, mais seulement quelques-uns d'entre eux (J=3/7 sous-tests, A=4/9 sous-tests).

Les résultats obtenus indiquent des différences significatives pour la plupart des comparaisons. Pour la version junior, des différences significatives sont obtenues pour les trois sous-tests. Les sous-tests *Barking* et *Simple RT* montrent de fortes tailles d'effet (d de Cohen = 1.11 et 1.67), alors que le sous-test *Balloon Hunt* montre une taille d'effet moyenne (0.74) alors que le résultat moyen obtenu par le groupe clinique reste dans la moyenne normative (cote standard moyenne de 9.0). Pour la version adolescent, le sous-test *Hector Cancellation* ne distingue pas les groupes, alors que les sous-tests *Vigil*, *Simple RT* et *RBBS* distinguent les deux groupes avec des tailles d'effet moyenne à forte (0.74, 1.13 et 0.87).

Une seconde étude clinique compare un groupe de 29 enfants ayant souffert d'un accident vasculaire cérébral à un groupe contrôle apparié. Chez les plus jeunes (J), des différences significatives sont obtenues aux trois sous-tests présentés. Ici, la taille d'effet obtenue au sous-test *Barking* est moyenne (0.62), le résultat moyen obtenu par le groupe clinique étant au niveau de la moyenne normative, alors que celles obtenues aux sous-tests *Balloon Hunt* et *Hide & Seek Visual* sont fortes (1.10, 0.91). Chez les plus vieux (A), seul un sous-test sur les trois administrés permet de distinguer les groupes (*Hecuba Visual Search* avec une taille d'effet très forte de 1.43). Il n'y a donc pas de différence significative entre les groupes pour les sous-tests *Hector Cancellation* et *Vigil*.

Enfin, ils ont comparé les résultats obtenus par un groupe de 24 enfants traités pour une leucémie lymphoblastique aigüe à un groupe contrôle. C'est la seule étude clinique qui présente des résultats pour les 9 sous-tests de la batterie ainsi que pour les scores globaux. Aucune différence significative n'a été trouvée entre les deux groupes.

Critique du test

Voici notre opinion préliminaire suite à l'analyse du contenu du test, une brève expérience d'administration clinique et les informations probantes disponibles à ce jour (en date de décembre 2018). Cette critique vise surtout à stimuler la réflexion des professionnels par rapport à l'interprétation clinique des résultats dans une démarche diagnostique auprès d'une population clinique. Chaque utilisateur est néanmoins invité à lire le manuel d'utilisateur attentivement afin de se forger son propre avis.

Points forts :

- Les nouvelles normes plus récentes sont basées sur un échantillon représentatif de la population du Royaume-Uni.

- La taille de l'échantillon normatif est beaucoup plus grande que celui de la version originale précédente.
- Les stimuli et l'exigence des tâches sont mieux adaptés à l'âge avec les deux versions offertes, ce qui est intéressant pour l'administration auprès d'une population clinique ayant des retards ou troubles quelconques (ex. moins de tâches impliquant un dénombrement avancé ou à rebours, moins de charge en mémoire de travail, moins longue durée des sous-tests, etc.).
- Un des rares tests à offrir des normes dès l'âge de 5 ans.
- Le protocole imprimé en couleur ainsi que les stimuli informatisés sont attirants.
- Une fiche de motivation est incluse pour les plus jeunes (un autocollant en forme d'étoile est posé sur un certificat pour chaque épreuve complétée).
- L'administration est simple et conviviale.
- Un chronomètre est inclus dans le logiciel pour les tâches papier-crayon avec un son de « départ » et un son « d'arrêt » ce qui limite les erreurs ou la variabilité d'administration.
- Le logiciel fonctionne sans la clé USB, ce qui permet l'utilisation sur plusieurs ordinateurs situés à des endroits physiques différents.

Points faibles :

- La cotation normative est entièrement informatisée. Le clinicien n'a donc aucun accès aux normes du test. Cette situation ne permet donc pas d'avoir un aperçu de la variabilité des scores en fonction de l'âge, ni même de la courbe développementale d'un âge à l'autre ou d'un sous-test à l'autre. Ceci réduit de façon considérable la flexibilité de l'interprétation des données (ex. on ne peut pas vérifier la cote qui aurait été obtenue à une tranche d'âge supérieure ou inférieure).
- Lorsque l'administration est terminée et que le résumé des scores en format pdf est demandé pour impression, le logiciel efface les données du client en question. Il n'est donc pas possible d'imprimer les résultats sans perdre les données. Ceci peut s'avérer problématique dans certaines situations (ex. après avoir imprimé des résultats, je constate une erreur d'entrée pour l'âge de l'enfant, rien à faire puisque les données sont effacées et que je n'ai aucun accès aux normes.).
- Le logiciel est long à charger sur ordinateur. La navigation d'une partie du logiciel à une autre n'est pas toujours aisée.
- Les paramètres de performance pris en considération dans la cotation sont très peu nombreux : un seul paramètre par sous-test est normé (p. ex., le temps de réaction, sans considération pour les erreurs).
 - Le test se démarque ainsi des autres tests populaires, tel que le CPT, le D-KEFS et la Nepsy-II, et surtout, de la tradition neuropsychologique, telle que l'approche qualitative proposée par Edith Kaplan ou les travaux de Donald Stuss sur les façons de mesurer les différentes composantes anatomo-fonctionnelles de l'attention. Les résultats normés au TEA-Ch2 permettent donc difficilement l'analyse de performance en vue de mieux qualifier les

fonctions en cause (s'agit-il d'un manque de contrôle attentionnel, d'un manque d'inhibition, d'une lenteur d'exécution, d'une combinaison de difficultés, etc.?). Par exemple, dans l'épreuve *SART* qui utilise un paradigme de « Go/No-go » similaire au CPT, le score au sous-test est uniquement déterminé par le nombre d'erreurs de commission, alors que le CPT-3 fournit neuf scores (temps de réaction, commission, omissions, persévération, variabilité, discrimination, etc.).

- Cet aspect réduit donc notablement l'efficacité de cette batterie à l'égard du temps investi puisque d'autres mesures seront nécessaires pour mieux qualifier les capacités cognitives attentionnelles et exécutives de l'enfant.
- Les auteurs du TEA-Ch2 n'offrent aucune justification théorique ou psychométrique à l'égard du choix du paramètre de performance unique retenu à chacun des sous-tests. De plus, les informations données sur le calcul des résultats sont parfois imprécises (ex. que se produit-il en cas de contradiction entre la réponse motrice et la réponse verbale aux sous-tests *Hide & Seek Auditory* ou *Cerberus* ?).
- Il est difficile de prendre en considération qualitative des mesures non normées aux sous-tests informatisés puisque le logiciel ne donne aucun autre résultat que celui qui est normé. Il est donc impossible de connaître, via le logiciel, le nombre de bonnes réponses ou le type d'erreurs commis aux sous-tests administrés à l'ordinateur. Il est donc essentiel d'observer l'enfant durant l'administration du test pour enregistrer des observations qualitatives supplémentaires. Ce type de considération qualitative doit toutefois être fait avec prudence puisque les données normatives aux autres tests d'attention sur le marché montrent bien que le taux d'erreur attendu change rapidement avec l'âge (surtout en bas âge).
- Les mesures sont majoritairement axées sur la vitesse d'exécution (7/9 sous-tests dans la version A). Par exemple, le nombre d'items identifiés aux tâches d'annulation équivaut à une vitesse de traitement puisque le temps est limité (plus l'enfant va rapidement, plus il identifie de cibles). De même, seul le temps de réaction est calculé au *Simple RT*, sans considération pour le nombre d'omission ou de commission. Ceci fait en sorte que les enfants impulsifs pourraient se voir avantagés par une vitesse d'exécution rapide, surtout compte tenu du fait que plusieurs sous-tests sont brefs. Au contraire, les enfants anxieux et minutieux pourraient se voir désavantagés même si leur performance est plus exacte.
- Malgré une mise à jour du logiciel, certains pépins sont remarqués, lesquels peuvent entraîner des résultats inattendus (ex. si un enfant ne réussit aucun item dans *RBBS*, il serait supposé n'avoir aucun résultat ou calcul. Or, dans nos essais, cette performance résulte en une très bonne performance en score standard). Des situations exceptionnelles restent donc à éclaircir, d'où la nécessité d'observer l'enfant durant l'administration du test pour noter ses erreurs et de suivre les mises à jour du logiciel, lorsque disponibles.

- Il est étonnant de constater que la batterie contient autant de sous-tests d'attention sélective visuelle (tâche d'annulation), alors que dans la première version du TEA-Ch, les deux épreuves d'attention visuelle sélective (*Sky search* et *Map mission*) étaient les moins discriminantes pour détecter les enfants ayant un TDAH, résultats qui ont aussi été obtenus par d'autres chercheurs (Heaton et al., 2001; Lemiere et al., 2010).
- Certaines caractéristiques des tâches semblent faciliter la performance pour les enfants ayant des difficultés attentionnelles et exécutives. Par exemple, le fait de pouvoir compter à voix haute à l'épreuve *Barking* ou *Vigil* diminue la charge d'attention et de mémoire de travail. De plus, la moitié des items sont à intervalle régulier, ce qui est plus facile. Autre exemple, la rétroaction donnée lors d'erreurs dans le sous-test *Simple RT* ou *RBBS* permet une rétroaction immédiate à l'enfant (ce qui est connu pour être aidant chez les enfants ayant un TDAH) et elle oblige l'enfant à se recentrer sur la tâche lors de moments d'inattention. Que mesure-t-on réellement lorsque l'attention de l'enfant est redirigée à la tâche? Enfin, la majorité des épreuves ont une durée brève (≤ 6 min.) la plus longue durant environ 10 minutes dans la version adolescent.
- Les indices psychométriques de certains sous-tests sont plus faibles et invitent à la prudence dans l'interprétation (fiabilité test-retest, validité convergente), particulièrement dans la version J.
- Les études cliniques sont limitées à plusieurs égards, notamment parce qu'aucune n'a été faite sur une clientèle ayant un TDAH confirmé et que les données ne sont présentées que pour un nombre restreint de sous-tests de la batterie. Nous avons donc très peu d'informations sur les indices réels de sensibilité et spécificité.

En résumé :

Nous sommes d'avis qu'il s'agit d'un test intéressant pour les enfants, facile à administrer et avec des normes récentes. Le format des activités proposées aux enfants permet aisément d'élargir notre champ d'observations qualitatives comportementales de l'enfant (ex. les lamentations face à l'effort soutenu ou à l'ennui, l'agitation motrice, l'anxiété, etc.). Par contre, pour le temps investi en administration, le clinicien en ressort avec peu de scores normés et peu de possibilités d'analyser les résultats pour qualifier les fonctions cognitives en cause dans les performances de l'enfant (ex. départager ce qui appartient à l'impulsivité, à l'inattention franche ou à la vitesse du traitement de l'information). L'administration informatique peut ainsi donner la fausse illusion que l'ordinateur note tout et se substitue aux observations qualitatives, alors qu'au contraire, les observations qualitatives détaillées faites durant la passation informatisée nous semblent essentielles pour guider l'interprétation clinique des résultats obtenus.

Par ailleurs, la sensibilité de cet outil chez les enfants présentant des difficultés cognitives d'attention n'est pas encore bien établie, même que la forme de certains sous-tests, le choix des scores utilisés dans la normalisation et les données psychométriques présentées nous permettent de supposer que la sensibilité de cette batterie puisse être limitée. D'ici la publication de données psychométriques supplémentaires sur la sensibilité et la spécificité de cette batterie, les cliniciens sont invités à la prudence dans leur interprétation clinique

des résultats normés obtenus à cette batterie. Rappelons que les données probantes nous indiquent qu'un enfant ayant un TDAH peut obtenir des résultats normaux à des tests psychométriques d'attention, de la même façon que des enfants n'ayant pas de TDAH puissent obtenir de faibles résultats à ce type de tests pour d'autres raisons cliniques qu'une problématique primaire au niveau attentionnel (Willcutt et al., 2005). Ainsi, les guides de pratique clinique connus en psychologie conservent toute leur valeur : le jugement clinique doit continuer de porter sur un ensemble d'informations, dont les observations comportementales faites durant l'administration, les résultats à d'autres mesures (ex. questionnaires) et les informations obtenues en entrevue (impact fonctionnel dans les différents milieux de vie, évolution des symptômes dans le temps, tableau de comorbidités, etc.).