



Master

2014

Open Access

This version of the publication is provided by the author(s) and made available in accordance with the copyright holder(s).

Le test d'apprentissage HART : l'effet d'évaluation dynamique sur des personnes avec et sans déficience intellectuelle

Zuretti, Chiara

How to cite

ZURETTI, Chiara. Le test d'apprentissage HART : l'effet d'évaluation dynamique sur des personnes avec et sans déficience intellectuelle. Master, 2014.

This publication URL: <https://archive-ouverte.unige.ch/unige:44145>



UNIVERSITÉ
DE GENÈVE

FACULTÉ DE PSYCHOLOGIE
ET DES SCIENCES DE L'ÉDUCATION

Le test d'apprentissage HART : l'effet de l'évaluation dynamique sur des personnes avec et sans déficience intellectuelle

**MEMOIRE REALISE EN VUE DE L'OBTENTION DU/DE LA
DE LA MAÎTRISE EN EDUCATION SPÉCIALE**

**PAR
CHIARA
ZURETTI**

DIRECTEUR DU MÉMOIRE

Marco Hessels

MEMBRES DU JURY

Mireille Bétrancourt

Hildalill Rojas

Genève, 30 juin 2014

**UNIVÉRSITÉ DE GENÈVE
FACULTÉ DE PSYCHOLOGIE ET DES SCIENCES DE L'EDUCATION
SCIENCES DE L'ÉDUCATION**

RESUME

Cette étude s'interroge sur l'impact d'un entraînement standardisé du raisonnement analogique lors du test HART, sur des personnes présentant ou non une déficience intellectuelle. Les principales données ont été recueillies par le biais de l'enregistrement des mouvements oculaires des participants, grâce auquel a été possible faire émerger des précisions relatives à la performance et aux principaux comportements de résolution mis en œuvre par les personnes. L'échantillon de base, sans déficience intellectuelle, a par la suite été comparé à un échantillon préexistant, contenant des personnes ayant le même âge mental que le premier mais ayant une déficience intellectuelle. Suite à un entraînement standardisé, les élèves ayant une déficience intellectuelle se montrent plus performants que ceux sans, qui sont non entraînés, et maintiennent la même évolution que les personnes sans déficience intellectuelle, entraînées. L'importance du premier temps d'encodage dans la résolution systématique et performante des tâches analogiques est également mis en évidence.



**UNIVERSITÉ
DE GENÈVE**

**FACULTÉ DE PSYCHOLOGIE
ET DES SCIENCES DE L'ÉDUCATION**

Le test d'apprentissage HART : l'effet de l'évaluation dynamique sur des personnes avec et sans déficience intellectuelle

Mémoire présenté pour l'obtention de la Maîtrise en Education Spéciale
par

Chiara Zuretti

Déposé en Juin 2014

Directeur du mémoire :
Dr. Marco G. P. Hessel

Membres du jury :
Prof. Mireille Bétrancourt

Hildalill Rojas

REMERCIEMENTS

...À tous les directeurs et directrices, aux enseignants, aux enseignantes et aux amis qui ont soutenu et collaboré à cette étude. Mais surtout aux enfants et aux adolescents qui y ont participé avec enthousiasme, sans lesquels cette recherche n'aurait pas pu avoir lieu.

À Marco Hessel, pour son aide dans la construction de ce travail et dans l'analyse statistique. Mais surtout à l'humanité dont il fait preuve et à sa positivité contagieuse.

À Sophie Schorno, pour sa permission de continuer sa recherche et pour son soutien constant, même dans les moments de découragement.

À Hildalill Rojas et à Mireille Bétrancourt, pour avoir accepté de faire partie de la commission de ce travail.

À Daniela Zambarbieri, à sa disponibilité et son intérêt pour cette étude ainsi que son savoir précieux et ses indications précieuses sur les mouvements oculaires.

À Silvia, à l'amitié, au soutien réciproque, aux fous rires et aux rencontres étranges dans la bibliothèque de la FPSE après 19h.

À Olli, à ses posters colorés qui ont mis de la couleur dans mes journées grises, à sa patience et à son courage lui ayant permis de tenir jusqu'au bout.

À Eva, à son inexplicable enthousiasme pour tout ce qui est compliqué.

À Sancho et Findus, à leur soutien félin, plein de ronronnements et de miaulements.

À Bimbo, pour sa compréhension de mon idée que nous ne sommes jamais trop vieux pour être fous.

À Tobii, fidèle et silencieux copain de voyage, parlant une langue compréhensible uniquement pour les extraterrestres et les informaticiens.

...À l'arbre vigoureux que je suis devenue grâce à vous.

À mon père, qui a toujours renforcé mon écorce, avec de nouveaux défis.

À ma mère, qui a toujours arrosé mes racines, sans rien me demander.

À Federica et à sa capacité de voir le vent souffler dans les branches de l'arbre et de prévoir la tempête.

À mes super grands-mères, à la terre riche et solide dans laquelle je plonge mes pieds.

Et à vous, César et Franco, rayons de soleil qui nourrissent chaque jour mes feuilles tournées vers le ciel.

Et à toi, ma petite Alba, ma fleur souriante, au bout de mes branches.

Et à la simple phrase que Claudia a certainement oublié m'avoir dite : « rien n'est impossible si tu y crois. »

RÉSUMÉ

Cette étude porte sur l'évaluation dynamique du potentiel d'apprentissage d'un individu. Elle s'inscrit dans l'analyse des effets d'un entraînement au cœur d'un test d'apprentissage, et plus précisément dans l'investigation de l'impact d'un entraînement standardisé du raisonnement analogique lors du test HART, sur des personnes présentant ou non une déficience intellectuelle. Les principales données de ce travail ont été recueillies par le biais de l'enregistrement des mouvements oculaires des participants, grâce auquel il a été possible de faire émerger des précisions relatives à la performance et aux principaux comportements de résolution mis en œuvre par les personnes, lors de la résolution d'une procédure du type pré-test – entraînement – post-test. Afin de saisir l'effet de l'entraînement sur les participants, il a été nécessaire de créer un groupe expérimental ainsi qu'un groupe de contrôle, pour l'échantillon de base, à savoir sans déficience intellectuelle. Ce dernier a par la suite été comparé à un échantillon préexistant, contenant des personnes ayant le même âge mental que le premier mais présentant une déficience intellectuelle. Les résultats montrent qu'il existe une différence entre les groupes au niveau de la performance. Suite à un entraînement standardisé, les élèves ayant une déficience intellectuelle se montrent plus performants que ceux sans, qui sont non entraînés, et ils maintiennent la même évolution que les personnes sans déficience intellectuelle, entraînées. Les changements de comportement lors de la résolution de l'analogie mettent en évidence l'importance du premier temps d'encodage dans la résolution systématique et performante des tâches analogiques.

Table des matières

Liste des tableaux	I
Liste des Figures	II
Introduction.....	III
I CADRE THEORIQUE	
1. Les Tests d'intelligence.....	1
1.1 <i>Contexte historique et social</i>	1
1.2 <i>Le concept de standardisation</i>	2
2. Les critiques des tests de QI.....	3
2.1 <i>Les bases théoriques</i>	3
2.2 <i>Stabilité des résultats et stabilité de l'intelligence</i>	5
2.3 <i>La performance : un produit</i>	7
3. Les Tests d'apprentissage	8
3.1 <i>L'évaluation dynamique</i>	8
3.2 <i>Fidélité et validité</i>	10
4. Le HART	11
4.1 <i>Caractéristiques de l'outil</i>	11
4.2 <i>La phase d'entraînement</i>	12
4.3 <i>Études de validation du HART</i>	14
5. Le raisonnement	16
5.1 <i>Le raisonnement inductif</i>	16
5.2 <i>Le raisonnement analogique</i>	16
5.3 <i>Tâches analogiques</i>	17
5.4 <i>Composantes en jeu</i>	18
5.5 <i>Études sur le raisonnement analogique</i>	19
6. Les mouvements oculaires.....	20
6.1 <i>Les yeux et leurs mouvements</i>	20
6.2 <i>Les fixations et les saccades</i>	21
6.3 <i>Études concernant les mouvements oculaires</i>	23
7. Question de recherche et hypothèses	28
II PARTIE EMPIRIQUE	
8. Méthodologie	30
8.1 <i>Échantillon</i>	30

8.1 Instruments	31
8.1.1 Les Matrices Progressives de Raven (1998)	31
8.1.2 Le Hessels Analogical Reasoning Test, version informatique	31
8.1.3 Eye trackers - Tobii Studio.....	32
8.2 Plan de la recherche.....	34
8.2.1 Passation du Raven SPM	34
8.2.2 Passation du HART	34
8.3 Procédure d'administration	35
8.3.1 Raven.....	35
8.3.2 HART	36
8.4 Méthode d'analyse des données	38
8.4.1 Les variables relatives aux scores des différents tests.....	38
8.4.2 Tobii studio.....	38
9. Présentation des résultats.....	42
9.1 Analyse de l'échantillon de base	42
9.1.1 Scores aux tests	42
9.1.2 Analyse des données de Tobii Studio.....	43
Encodage.....	44
Matrice.....	44
Alternatives	44
Ailleurs	44
Temps total.....	46
Comparaisons entre matrice et alternatives	46
9.2 Analyse des deux échantillons, avec et sans déficience intellectuelle	48
9.2.1 Scores aux tests	48
9.2.2 Analyse des données de Tobii Studio.....	48
Encodage.....	49
Alternatives	49
Matrice.....	50
Ailleurs	50
Temps total.....	50
Comparaisons entre matrice et alternatives	51

10. Discussion	53
11. Limites et perspectives nouvelles	57
<i>11.1 Limites</i>	57
<i>11.2 Perspectives</i>	58
12. Conclusion	59

VI BIBLIOGRAPHIE
V ANNEXES

<i>Annexe A : Protocole de passation du test Raven SPM (1998)</i>	71
<i>Annexe B : Protocole de passation du test HART, version française</i>	75
<i>Annexe C : Protocole de passation du test HART, version italienne</i>	85
<i>Annexe D : Items utilisés pour le pré-test du HART</i>	92
<i>Annexe E : Exercices et items utilisés pour l'entraînement</i>	96
<i>Annexe F : Items utilisés pour le post-test</i>	99
<i>Annexe G : Exemples d'images découlant de l'outil Gaze Plot</i>	103

Liste des tableaux

<i>Tableau 1. Caractéristiques de l'échantillon final.....</i>	30
<i>Tableau 2: Phases de passation de la recherche.....</i>	35
<i>Tableau 3. Moyennes, Ecarts-types, F, df, p, aux pré- et post-tests du HART des 37 élèves</i>	42
<i>Tableau 4. Corrélation de Pearson entre Raven SPM, âge chronologique, Hart pré et post-test pour GC et GE.....</i>	43
<i>Tableau 5. Moyennes et Écarts-types de l'échantillon final pour le HART et données statistiques pour chacune des variables</i>	45
<i>Tableau 6:Corrélations de Pearson (signification unilatérale) entre pré- et post-test avec données Tobii pour GCo</i>	46
<i>Tableau 7:Corrélations de Pearson (signification unilatérale) entre pré- et post-test avec données Tobii pour GCo</i>	47

Liste des figures

<i>Figure 1</i>	<i>Item A1</i>	32
<i>Figure 2</i>	<i>Système d'enregistrement des mouvements oculaires</i>	33
<i>Figure 3</i>	<i>Évolution du score au HART pour GCd, GEd GCo, GEo</i>	49

Introduction

Ce travail s'inscrit dans la conception de l'intelligence soutenue par l'éducation cognitive, selon laquelle l'intelligence est considérée comme la *capacité à apprendre*, un potentiel modifiable qui se développe selon le contexte et les opportunités de chaque personne (Carlson & Wiedl, 1992/1995; Paour & Cèbe, 1999).

À la différence de l'évaluation traditionnelle des capacités intellectuelles, qui se basent sur ce que la personne a appris jusqu'au moment de l'examination de ses capacités, l'évaluation dynamique vise à comprendre ce qu'elle peut encore apprendre. En d'autres termes, elle s'intéresse au potentiel d'apprentissage. Les tests d'apprentissage offrent la possibilité d'évaluer ce dernier, en transmettant à l'individu les compétences nécessaires pour la résolution des items du test. Nombre de chercheurs ont confirmé les bénéfices de l'évaluation dynamique pour la mesure des capacités des personnes, aussi bien dans le cas de populations dites « à risque » (comme par exemple Hessels, 1997 ; Klauer, 1998 ; Schorno, 2013 ; Schlatter, 1999 ; Wishart & Duffy, 1990) que dans celui d'individus ordinaires (voir par exemple Berger, Bosson & Hessels, 2004 ; Rumley, 2007). Grâce aux études des dernières décennies, certains auteurs (Bethge, Carlson et Wiedl, 1982 ; Hessels, Vanderlinden, & Rojas, 2011 ; Bethell-Fox, Lohman & Snow, 1984 ; Rayner, 1998 ; Vigneau, Caissie & Bors, 2006) ont pu mettre en évidence la richesse des informations obtenues par le biais de l'enregistrement des mouvements oculaires, qui permettent de mieux comprendre les processus de raisonnement sous-jacents à la résolution.

Cette étude vise à analyser les changements de performance et de comportement lors de la résolution d'un test d'apprentissage analogique chez des élèves ayant le même âge mental, avec et sans déficience intellectuelle. Pour ce faire, il a été nécessaire de diviser l'échantillon de base en deux groupes. Le premier groupe, de contrôle, a suivi une procédure du type « pré-test-post-test » ; le deuxième, le groupe expérimental, a quant à lui bénéficié d'une phase d'entraînement entre le pré- et le post-test. Les données concernant les mouvements oculaires ont été enregistrées pendant les phases de pré- et de post-test, et ce pour les deux groupes. L'échantillon de base a par la suite été comparé à un échantillon du même âge mental, présentant une déficience intellectuelle et provenant d'une étude récente, centrée sur le même test d'apprentissage (Schorno 2013).

Ce travail se structure en trois parties. Premièrement, le *Cadre théorique* illustrera le contexte historique et social ayant mené à la naissance des tests traditionnels d'intelligence. Les principales critiques de ces outils ainsi que la naissance de l'évaluation dynamique et les avantages de l'utilisation des tests d'apprentissage, seront ensuite passés en revue. Un chapitre en particulier sera consacré à la présentation du test d'apprentissage utilisé pour cette recherche, le HART, et aux principales études qui valident son utilisation auprès de différents types de population. Le raisonnement analogique, impliqué dans les tâches de nombreux tests, qu'ils soient classiques ou dynamiques, sera présenté. Un accent sera mis sur les processus sous-jacents à sa mise en œuvre et aux études qui en attestent l'efficacité. Une dernière partie du *Cadre théorique* se focalisera sur la compréhension des mouvements oculaires, notamment sur les mouvements de saccades et sur les fixations, pour passer dans un second temps à la présentation des principales études réalisées autour de l'enregistrement des mouvements oculaires pendant la lecture, mais surtout lors de la passation des tâches de *problem solving*, typiquement adoptées dans les tests d'intelligence.

Deuxièmement, les hypothèses et la question de recherche seront formulées, pour ensuite en venir à la *Partie empirique*. La méthodologie de ce travail sera alors expliquée, en mettant l'accent sur les divers instruments utilisés, les procédures de passation et la présentation des échantillons. Les résultats de l'étude seront illustrés à l'aide de tableaux résumant les éléments principaux.

La troisième partie, intitulée *Discussion et conclusion*, articulera une discussion qui aura pour but de répondre aux hypothèses et à la question de recherche, à l'aide de la théorie mobilisée pour cette étude. Un dernier chapitre se focalisera sur les limites de ce travail et sur les perspectives futures des études à venir.

I CADRE THEORIQUE

1. Les Tests d'intelligence

1.1 *Contexte historique et social*

Au cours du 20^{ème} siècle, l'évaluation de l'intelligence acquiert une importance croissante pour les scientifiques et les différents États de l'Europe occidentale. La raison principale justifiant la naissance de « dispositifs d'observation standardisés » (Huteau & Lautrey, 1999, p. 1) dérive d'une nécessité sociale : les tests sont la réponse à un besoin pratique des domaines de l'éducation, du travail et de la santé (Huteau & Lautrey, 2006). Ce besoin est influencé par le contexte historique: à la fin du 19^{ème} siècle, ces pays vivent un important changement culturel, social et économique. L'industrialisation de la société nécessite une généralisation de l'instruction (Chartier & Loarer, 2008), rendant ainsi indispensable la naissance d'une nouvelle évaluation des capacités des personnes, afin de pouvoir les sélectionner, les distinguer et les classer, aussi bien à l'école, que dans les usines et à l'armée (Huteau & Lautrey, 1999). Les premiers tests d'intelligence se développent dans les années 1890 et c'est au début du 20^{ème} siècle qu'ils deviennent un outil central, souvent le seul, utilisé par les psychologues pour l'évaluation des capacités intellectuelles (Huteau & Lautrey, 1999). Avec la création de l'école publique obligatoire, les tests deviennent la méthode utilisée pour distinguer les capacités des enfants. En France, suite à la mise en œuvre des lois Ferry de 1881 et 1882, les enfants provenant de milieux différents et présentant des capacités vivement hétérogènes sont soumis à l'obligation scolaire (Chartier & Loarer, 2008). C'est dans ce contexte qu'une commission ministérielle demande à Binet de construire un test permettant de distinguer les enfants capables de suivre un parcours scolaire classique de ceux étant incapables de le faire, et par conséquent destinés à être accueillis dans des institutions spéciales (Huteau & Lautrey, 2006). L'échelle métrique de l'intelligence créée par Binet et Simon en 1905 est désormais considérée comme le premier test d'intelligence ayant inspiré les scientifiques de l'époque pour la construction et l'adaptation d'instruments futurs (Bernaud, 2009). À l'époque, Binet se montre déjà très critique par rapport à l'évaluation des capacités cognitives à travers des processus élémentaires, inspirés de pratiques de laboratoire (Chartier & Loarer, 2008). Il met en avant la complexité de l'intelligence et souligne l'importance « d'étudier des fonctions intellectuelles plus complexes » (Bernaud, 2009, p. 12). Son but est d'évaluer l'intelligence sans pour autant impliquer les apprentissages antérieurs de la personne (Hessels & Hessels-Schlatter, 2010). L'échelle créée par Binet et

Simon permet de définir les enfants à travers leur âge chronologique et leur âge mental, correspondant à la performance obtenue dans le test (Chartier & Loarer, 2008). Les principes à la base de cet outil se perdent malheureusement au cours du temps, suite à sa traduction anglaise, utilisée aux États-Unis (Hessels & Hessels-Schlatter, 2010). Influencée par les courants de l'époque, qui la conçoivent comme comportant des caractères héréditaires et immuables, la mesure de l'intelligence est rapidement interprétée comme le Quotient Intellectuel (Hessels & Hessels-Schlatter, 2010). Cette définition, proposée en 1912 par Stern, consiste à définir les capacités intellectuelles sur la base d'un rapport entre l'âge mental et l'âge chronologique de la personne (Chartier & Loarer, 2008).

Chaque pays Européen voit ces instruments évoluer singulièrement. Il est cependant possible de relever de manière générale des tendances qui font surface sur la majorité du continent. Deux années importantes sont retenues par rapport à ces tendances: 1920, qui signe le début d'un « véritable mouvement des tests » (Huteau & Lautrey, 1999, p. 1) et 1960, période pendant laquelle l'efficacité de ces procédures est remise en question. En effet, certains scientifiques de la psychologie cognitive mettent en doute l'origine et la nature de l'intelligence ainsi que la modalité grâce à laquelle cette dernière est évaluée. Suite à ces premières critiques, de nouveaux outils d'évaluation se développent.

Dès leur création, les évaluations de l'intelligence prennent différentes formes, suite aux critiques posées au fil du temps et à l'influence des nécessités pratiques relatives à chaque époque (Huteau & Lautrey, 1999). Aujourd'hui, les tests de QI demeurent grandement exploités dans le cadre des écoles obligatoires: ils sont effectivement utilisés pour évaluer les capacités d'un enfant qui, pour différentes raisons, présente des difficultés d'apprentissage. Un résultat faible à ce type de test peut être utilisé comme un indicateur de la nécessité de placer l'enfant dans une classe spécialisée (Budoff, 1987a).

1.2 Le concept de standardisation

Le contenu ainsi que les évaluations des tests de QI sont variés et mesurent différents domaines de l'intelligence en faisant recours à plusieurs types d'exercices (Huteau & Lautrey, 1999). Malgré ces différences, ils répondent tous à un critère commun : la standardisation. Cette dernière caractérise principalement trois éléments du test, à savoir le matériel utilisé, la modalité de passation (présentation, consignes, situation d'observation, intervention de l'examinateur) et

l'évaluation de la personne concernée (Huteau & Lautrey, 1999). Chaque passation doit donc avoir les mêmes caractéristiques, égales pour toutes les personnes évaluées, en permettant ainsi de reproduire le même type de situation d'estimation pour tous. Grâce à cet aspect, il est possible de maîtriser certaines variables pouvant influencer l'évaluation de la personne, comme par exemple le comportement subjectif de l'examineur dans la phase du test. En outre, la standardisation permet de comparer la performance individuelle d'un élève à la moyenne des individus correspondant à son groupe de référence (Huteau & Lautrey, 1999). Si la standardisation n'est pas assurée, il est impossible de comparer les différentes personnes, car les situations d'évaluation ne répondent plus à des critères objectifs (Chartier & Loarer, 2008).

2. Les critiques des tests de QI

Au cours de l'histoire, les tests de QI sont soumis à différentes critiques. À partir de contestations plus générales concernant leur fonctionnement, des remarques relatives à l'inadéquation de ces outils pour l'évaluation des populations dites « à risque » seront présentées. Ce terme, « à risque », utilisé par Hessels (1997), regroupe différentes typologies d'individus, aux caractéristiques variées : les enfants impulsifs (Wiedl, 1990), les enfants des minorités ethniques (Hessels, 1997), les enfants avec une déficience intellectuelle ou une difficulté d'apprentissage (Hessels-Schlatter & Hessels, 2009a) et les personnes présentant un retard mental modéré à sévère (Hessels-Schlatter & Hessels, 2009b).

2.1 Les bases théoriques

Une première remarque vise à souligner un manque de bases théoriques solides dans les tests de QI (Hessels & Hessels-Schlatter, 2010), car comme le soulignent Huteau et Lautrey (1999) « beaucoup de tests ont été construits sur une base empirique et ont donc des fondements théoriques faibles, ce qui conduit à une certaine incertitude quant à la nature des processus effectivement évalués » (p. 3).

Un accent est mis sur la validité prédictive de ces outils, car la corrélation entre la réussite scolaire et la performance aux tests n'est pas satisfaisante : elle varie entre .20 et .60, prédisant « qu'à peine la moitié de la variance de la réussite scolaire, et dans certains cas beaucoup moins » (Hessels & Hessels-Schlatter, 2010, p.7). Il est donc certain que les tests classiques

mesurent autre chose que la réussite scolaire et que, dans le cas d'enfants « à risque » ces facteurs influent davantage, résultant en une corrélation d'autant plus faible (Hessels & Schlatter, 2010).

Cet aspect est à mettre en relation avec l'incohérence entre la conception de l'intelligence et son opérationnalisation effective (Beckmann, 2006 ; Hessels *et al.*, 2008). En effet, comme l'explique Beckmann (2006), l'intelligence est définie comme « the ability to learn » (p. 37), alors que son estimation ne donne pas la possibilité d'évaluer ce que l'enfant peut réellement apprendre. Le résultat à une épreuve classique ne peut donc pas être indicatif des capacités d'apprentissage de la personne, car aucun apprentissage n'est inclus dans l'évaluation (Hessels & Hessels-Schlatter, 2010). À cause de cela, au lieu de mettre en évidence la capacité d'apprendre de l'enfant, les tests de QI se limitent à fournir une estimation de ce qu'il a appris jusqu'au moment de l'évaluation (Hessels, Vanderlinden, & Rojas, 2011) en mesurant des connaissances préscolaires et des apprentissages antérieurs qui varient d'un enfant à l'autre et qui ne représentent pas leur réelle capacité d'apprentissage (Berger, Bosson et Hessels, 2004).

À ce fait s'ajoute l'idée que les évaluations classiques se basent sur le préconcept que tous les enfants bénéficient des mêmes possibilités d'apprentissage, alors qu'en réalité l'accès aux connaissances est très varié, dépendant des possibilités et des stimuli auxquels l'enfant est confronté dans son milieu familial ainsi que de la culture à laquelle il appartient (Bernaud, 2009 ; Budoff 1987a ; Campione & Brown, 1987 ; Hessels & Hessels-Schlatter, 2010). En opérant de telle manière, les tests de QI comparent une population très hétérogène en supposant que cette dernière a eu les mêmes possibilités d'apprentissage, ce qui est bien sûr erroné, car, comme le soulignait déjà Rey :

[...] chez les uns, les possibilités d'adaptation se sont actualisées au maximum, chez d'autres, des marges plus ou moins étendues subsistent encore ; chez d'autres enfin, les accommodats ne se sont pas encore constitués, ce qui n'implique nullement qu'ils ne puissent apparaître un jour. (Rey, 1934, p. 299)

Comme l'explique Bernaud (2009) « l'intelligence est un construit culturel, car elle s'applique sur des contenus (langagiers ou autres) et des procédures de résolution des problèmes intimement liés à la culture » (p. 26). Pour Budoff (1987a) l'intelligence ne peut être détachée de la culture à laquelle appartient l'individu évalué. Si les tests classiques mesurent les capacités antérieures des

personnes, en se basant sur l'égalité des chances, il est évident que ces outils sous-estiment tout un ensemble de personnes qui n'appartient pas à la culture et à la classe sociale ayant construit les tests, car « plutôt que d'évaluer un construit scientifique, [ils] mesureraient un savoir social réservé à l'élite, contribuant de la sorte à renforcer la discrimination et l'exclusion sociale » (Bernaud, 2009, p. 18). Il s'agit ici d'enfants d'origines sociales défavorisées et de ceux issus de minorités ethniques (Budoff, 1987a ; Campione & Brown, 1987). Hessels (1997), dans une étude relative à ces derniers, démontre que leur résultat à un test de QI peut donner une mauvaise estimation de leurs capacités cognitives. Pour ces enfants, la passation à un test traditionnel peut les confronter avec une évaluation directe des capacités linguistiques et culturelles, ou avec une évaluation dans laquelle le côté verbal n'est pas directement évalué mais tout de même influant sur les performances (Hessels, 1997). Budoff (1987a) explique que ces enfants se démontrent capables de résoudre des problèmes et d'acquérir des connaissances importantes pour leurs communautés d'appartenance : ils ont donc des capacités d'apprentissage que les tests classiques ne peuvent relever. Ce fait s'explique par une impossibilité d'évaluer, avec le même outil, des individus qui voient leur intelligence se développer selon les nécessités imposées par un contexte différent. De plus, cette discrimination ne se limite pas à des problèmes de nature culturelle et sociale. En effet, les tests classiques se révèlent également être inadaptés pour les personnes présentant des difficultés intellectuelles. En ce qui concerne les individus ayant une déficience intellectuelle modérée à sévère, Hessels-Schlatter (2010) souligne que « les déficiences motrices, sensorielles et langagières entravent fortement leurs performances dans les tests traditionnels » (p. 35). En outre, ces personnes se démontrent souvent incapables d'accéder au raisonnement analogique, élément nécessaire pour la résolution de la plupart des tests de QI, étant donné que les processus cognitifs impliqués dans ce dernier sont souvent déficitaires (Hessels-Schlatter & Hessels, 2009b).

2.2 Stabilité des résultats et stabilité de l'intelligence

Une deuxième critique soulevée met l'accent sur le caractère stable des résultats aux tests classiques, avec lesquels il est impossible de concevoir une évolution des capacités cognitives de la personne. Ces outils évaluent les compétences intellectuelles des individus en se basant sur une épreuve unique, qui ne permet pas « [d'] obtenir la meilleure performance de l'individu testé » (Hessels & Hessels-Schlatter, 2010, p.9) car plusieurs facteurs non cognitifs peuvent

influencer le résultat. Selon Campione & Brown (1987), ce type de procédure évalue l'enfant avec un chiffre représentant ses capacités à un moment précis de son existence, et qui, seul, explique ses capacités dans toutes situations et circonstances. Ce chiffre reflète donc son résultat de manière rigide, rendant ainsi impossible d'envisager une éventuelle amélioration (Campione & Brown, 1987). La stabilité est un aspect particulièrement sensible dans le cas d'une personne à la performance faible: le résultat à ce type de test risque de l'emprisonner dans une évaluation figée, influençant également l'enseignant, qui pourrait ne plus s'investir dans l'apprentissage de l'enfant et perdre l'espoir d'un véritable progrès. Cet aspect est surtout délicat pour les personnes présentant une déficience intellectuelle, car en n'étant pas habituées à des situations d'évaluation ni à la maîtrise du matériel proposé, elles ne peuvent souvent pas être évaluées pour ce qu'elles sont réellement capables de faire (Hessels-Schlatter, 2010). Dans la modalité classique de passation du test, la personne est censée résoudre les problèmes présentés sans recevoir d'explications précises sur la façon d'agir. Il n'y a donc pas de certitude que le sujet ait compris ce qu'il doit faire (Campione & Brown, 1987). Hessels-Schlatter et Hessels (2009a) soulignent que les enfants ayant des difficultés d'apprentissage ou une déficience intellectuelle sont souvent sous-estimés par les tests de QI, car ils n'arrivent pas à comprendre ce que nous leur demandons de faire ni comment le faire. Cela résulte en un manque d'engagement de leur part et une production de réponses au hasard. Ce type de comportement est également souligné chez les jeunes enfants (Hessels & Tieksstra, 2010). Une étude de Wishart et Duffy (1990) sur les enfants trisomiques confirme la faible fidélité des tests d'intelligence pour ce type de population. Les auteures démontrent que, malgré le fait que la performance globale des participants ne varie pas significativement, les *patterns* de réponse dans les deux sessions de l'évaluation sont instables : les enfants échouent dans la résolution des items réussis dans la session précédente puisqu'ils ne sont pas engagés dans les tâches proposées. Comme le souligne Haywood (1977), le fait qu'un enfant n'arrive pas à faire un exercice ne signifie pas qu'il n'en ait pas les capacités. Malgré cela, souvent, le faible score de ces enfants est généralement associé à une capacité d'apprendre plus mince (Budoff, 1987a ; Hessels-Schlatter & Hessels, 2009a).

La relation causale entre le niveau de performance et les capacités d'apprentissage d'une personne, où la deuxième serait la conséquence directe de la première, découle d'une conception particulière de l'intelligence soutenue par la majorité des inventeurs des tests classiques. En

effet, l'intelligence a longtemps été conçue comme un élément stable et déterminé. Carlson et Wiedl (1992/1995) mettent en évidence une différence dans la représentation de l'intelligence, découlant de deux diverses approches pour l'évaluation de cette dernière. D'un côté, « les théoriciens de la stabilité » (p. 217), conçoivent l'intelligence humaine comme un élément figé, imperméable au changement et immuable dans le temps. Cette idée porte à concevoir les capacités de l'individu de manière rigide. Comme l'explique Büchel (2007): « [l']intelligence est considérée dans cette approche comme un trait de la personnalité déterminant les limites du développement cognitif » (p. 35), influençant par conséquent l'idée d'une possible amélioration. Cependant, plusieurs auteurs (voir par exemple Budoff 1987a ; Büchel, 2007 ; Feuerstein 1979 ; Haywood, 1977; Hessels & Schlatter, 2010) nous invitent à considérer l'intelligence comme possédant des caractéristiques dynamiques. Cette vision, soutenue par « les partisans du changement » (Carlson & Wiedl, 1992/1995, p. 216), s'oppose à la première et considère les capacités de toute personne comme étant modifiables dans le temps, en interaction avec l'environnement et influencées par des facteurs externes. Cette deuxième représentation de l'intelligence caractérise la théorie sur laquelle se basent les tests d'apprentissage et conçoit l'intelligence comme la *capacité à apprendre*, en se focalisant ainsi sur le potentiel de chaque individu (Paour & Cèbe, 1999).

2.3 La performance : un produit

Une troisième remarque porte sur la rigidité de l'évaluation des tests traditionnels qui expliquent la performance de l'élève en se basant uniquement sur son produit final. Ces outils ne donnent aucune information sur les processus mis en place pour arriver au résultat (Campione & Brown, 1987 ; Feuerstein, 1979) et considèrent le produit comme synonyme de la performance de la personne évaluée. Beckman (2006) met l'accent sur l'incompatibilité des évaluations classiques puisqu'elles considèrent le produit final comme étant représentatif des processus, en sous-estimant celui de la personne évaluée. Pour Boudoff (1987a), les différences de performance aux tests de QI sont mal interprétées puisqu'elles se limitent à expliquer les faibles résultats comme étant représentatifs d'une intelligence moindre. Généralement, les tests classiques sont « massivement inspirés par une approche centrée sur les produits » (Bernaud, 2009, p. 17). S'intéresser au cheminement et aux processus utilisés lors de la résolution du test permettrait d'aller plus loin dans leur compréhension, en évitant de se limiter à l'explication de la

performance à travers la réussite ou l'échec (Bernaud, 2009), ce qui entrave tout renseignement sur des pistes éducatives à suivre pour améliorer les capacités de l'individu et sur ce qu'il pourrait encore apprendre (Hessels & Hessels-Schlatter, 2008).

Comme nous l'avons vu dans ce chapitre, les tests classiques de QI comportent plusieurs défauts, aussi bien dans leur construction théorique que dans leur mise en pratique effective. Ce type d'évaluation est spécialement dévalorisante pour les enfants « à risque », pour qui une procédure plus respectueuse de leurs conditions serait souhaitable. Celle-ci permettrait « d'assurer la compréhension des instructions, de clarifier les attentes et de réduire les biais dus à des difficultés de langage et à la culture » (Hessels & Hessels-Schlatter, 2010, p. 10).

3. Les Tests d'apprentissage

3.1 L'évaluation dynamique

À partir de ces différents constats, au début des années 1960, certains chercheurs ont développé une alternative à l'évaluation traditionnelle de l'intelligence, ayant pour but de donner une estimation plus valable de la capacité d'apprentissage et une meilleure prédiction pour les enfants « à risque » que celle fournie par les tests de QI (Beckmann, 2006 ; Hessels, Berger, Bosson, 2008). Il s'agit des tests d'apprentissage, qui se caractérisent par une évaluation dynamique.

La principale différence entre les tests de QI et les tests d'apprentissage consiste en le choix des critères d'évaluation. Les évaluations psychométriques classiques se basent sur le concept de stabilité. Ce dernier concerne le trait que nous choisissons de mesurer et les résultats obtenus (Hessels-Schlatter & Hessels, 2009a). En effet, pour que les données puissent être valables, aucun type de relation avec l'examinateur n'est prévu pendant la résolution du test, car ceci est considéré comme un facteur nuisant à la fidélité des résultats. Afin que les données des différents participants puissent être comparées, l'interaction entre l'administrateur et l'enfant doit pourtant être minimale (Hessels-Schlatter & Hessels, 2009a). L'intervention de l'administrateur est dite statique, car elle est généralement limitée à des échanges ayant comme but principal celui de présenter les exercices et les questions auxquelles l'enfant doit répondre, tout en maintenant une

bonne relation avec lui (Hessels-Schlatter & Hessels, 2009a). À la différence de la notion de stabilité, les tests d'apprentissage se caractérisent par une évaluation dynamique, incluant une phase d'apprentissage (Hessels-Schlatter & Hessels, 2009a). Cette dernière est présentée aux participants sous forme d'entraînement standardisé, modifiant ainsi le rôle de l'examinateur durant la passation. Contrairement aux évaluations classiques, l'examinateur entre en réelle interaction avec l'élève, en lui enseignant les éléments nécessaires à la résolution des items (Hessels-Schlatter & Hessels, 2009a). La phase d'apprentissage est fondamentale, car elle permet de réduire les inégalités d'apprentissages antérieurs entre les enfants et vise à évaluer ce que ces derniers peuvent véritablement apprendre, et non pas ce qu'ils ont déjà acquis (Budoff, 1987a). Grâce à ce changement, les tests d'apprentissage peuvent être considérés comme étant plus représentatifs de l'évaluation de l'intelligence, puisqu'ils permettent de mesurer les capacités de l'enfant dans une réelle situation d'apprentissage (Hessels & Tiekstra, 2010).

Plusieurs critiques ont néanmoins été soulevées pour ce type d'évaluation. La principale remarque consiste à soutenir l'idée que l'évaluation dynamique n'est pas une procédure d'évaluation standardisée, en focalisant la critique sur le rôle actif de l'examinateur, considéré comme un obstacle à la fidélité de l'instrument. Comme l'explique Beckmann (2006) le concept d'évaluation dynamique est un terme large, qui inclut des approches aux formes et buts différents. En effet, ce même mot est représentatif de divers types d'évaluations, aux objectifs différents. D'un côté, l'évaluation dynamique est utilisée lors d'interventions cliniques développées par Feuerstein (1979), qui ne respectent pas le critère de standardisation : elle vise la construction d'un bilan qualitatif et individuel, permettant d'obtenir des pistes pédagogiques à suivre afin d'améliorer les capacités de la personne (Chartier & Loarer, 2008). L'interaction de l'examinateur est subjective, car il intervient pour donner une aide spécifique et particulière aux besoins de l'enfant (Dias, 2001). D'un autre côté, l'évaluation dynamique caractérise également les tests d'apprentissage. Ces outils, développés pendant les dernières décennies, répondent à des critères psychométriques de standardisation précis et permettent d'obtenir des données fidèles et valides (Beckmann, 2006). Cette confusion conceptuelle mélange donc les caractéristiques de différents outils et favorise, encore aujourd'hui, la fausse conception que les tests d'apprentissage ne sont pas des outils standardisés (Hessels-Schlatter & Hessels, 2009a).

3.2 Fidélité et validité

À la différence des instruments classiques, les tests d'apprentissage se montrent plus fidèles et valides dans l'évaluation des capacités cognitives pour les populations à risque et sont moins influencés par des facteurs environnementaux (Berger *et al.*, 2004). De plus, les tests d'apprentissage réduisent l'influence de facteurs non cognitifs, comme l'anxiété et l'hyperactivité, qui dans les tests classiques, peuvent suggestionner la performance et biaiser les résultats (Hessels *et al.*, 2008).

Suite au constat d'une surreprésentation des enfants provenant de minorités ethniques dans les écoles spécialisées aux États-Unis, Budoff (1987b) développe plusieurs outils d'évaluation du potentiel d'apprentissage en se basant principalement sur des exercices issus des tests de QI, comme par exemple le Block Design Test de Kohs. Budoff et Friedman (1964) utilisent cet instrument pour évaluer des adolescents catégorisés comme « mentally retarded » (p. 435) par une évaluation classique. La modalité de passation prévoit un pré-test, un entraînement lors duquel les personnes reçoivent les connaissances nécessaires pour la résolution des tâches, et un post-test. À travers ses résultats, Budoff (1967) met en évidence trois catégories de personnes : les *high scorers*, qui réalisent déjà une belle performance au pré-test ; les *gainers*, qui n'obtiennent pas un score élevé au pré-test mais qui améliorent significativement leur performance suite à l'entraînement, et les *non-gainers*, pour lesquels aucune amélioration n'est perçue : les résultats restent faibles avant et après la phase d'entraînement. Avec cette expérience, Budoff souligne l'inadaptation des tests de QI pour l'évaluation de personnes présentant une déficience intellectuelle, car certains d'entre eux se démontrent capables de profiter de l'expérience. Selon l'auteur, les tests de QI n'ont pas été créés pour concevoir cette distinction, renforçant ainsi la nécessité d'une évaluation dynamique de l'apprentissage pour cette population.

D'après Hessels (2009), les tests d'apprentissage offrent non seulement une estimation plus précise des capacités d'apprentissage des populations « à risque », mais permettent aussi une meilleure prédition de l'apprentissage dans d'autres domaines. Beckmann (2006) souligne la haute validité prédictive de la réussite scolaire des tests d'apprentissage par rapport aux tests de QI :

The direct comparison of both test approaches in a regression analytic design reveals, however, that the dynamic test approach is able not only to provide the same information

as traditional tests, but also additional information uniquely attributable to learning tests (Beckmann, 2006, p. 47).

Cet aspect est bien visible dans l'étude d'Hessels (1997). En utilisant le test LEM (Leertest voor Etnische Minderheden) avec des enfants néerlandais, turcs et marocains, il a pu démontrer que le LEM prédisait le même pourcentage de réussite scolaire qu'un test classique, et qu'à celui-ci s'ajoutaient 19% que le test de QI ne relevait pas. Ce test a par ailleurs été validé dans une autre étude en Suisse romande, démontrant des résultats similaires (Hessels & Hessels-Schlatter, 2010).

Bien que les tests d'apprentissage permettent de donner une meilleure estimation des capacités intellectuelles des populations à risque, leur utilisation est encore limitée dans la pratique, surtout à cause d'un manque de familiarisation de la part des professionnels (Hessels *et al.*, 2008). Il existe aussi d'autres facteurs qui ne facilitent pas leur mise en pratique. Premièrement, la tranche d'âge est souvent très limitée par rapport à l'étendue de celle proposée par des tests de QI (Hessels *et al.*, 2008). À cela s'ajoute le fait que, dans la majorité des cas, l'administration des tests d'apprentissage est individuelle et demande ainsi plus de temps pour sa mise en œuvre (Berger *et al.*, 2004).

4. Le HART

Sur la base des différentes limites mentionnées dans le chapitre précédent, Hessels élabora le Hessels Analogical Reasoning Test (HART), un test d'apprentissage construit tout en essayant de répondre aux principaux défauts de ce type d'instrument.

4.1 Caractéristiques de l'outil

Le HART est un outil d'évaluation du potentiel d'apprentissage dans le domaine du raisonnement analogique. Il a été construit pour des élèves âgés de 5 à 18 ans, provenant de l'enseignement ordinaire ou spécialisé (Hessels, 2009). Il s'agit d'un outil standardisé qui prévoit une évaluation dynamique. Il permet d'obtenir, dans un temps maximal de 60 minutes (Berger *et al.*, 2004), le niveau actuel du fonctionnement intellectuel ainsi que les possibilités d'apprentissage d'un enfant. La passation du HART peut être individuelle ou collective (Berger *et al.*, 2004).

Pour éviter des biais culturels et la confrontation de l'enfant à des connaissances linguistiques qui pourraient interférer avec sa compréhension, le test se base sur des protocoles géométrico-figuratifs (Hessels *et al.*, 2008). De plus, afin de ne pas faire émerger de sentiments négatifs liés à l'exécution de tâches scolaires dans lesquelles l'enfant pourrait se trouver en échec, les items proposés sont du type non-scolaire (Hessels & Tiekstra, 2010). En d'autres termes, ils représentent des exercices qui ne se focalisent pas sur la résolution de tâches contenant des notions spécifiques, mais plutôt sur des exercices qui favorisent l'utilisation de processus généraux, transférables dans plusieurs domaines. (Hessels *et al.*, 2008).

Les items du HART sont composés de dessins géométriques représentés sur des grilles de 2x3 ou 3x3, avec respectivement six ou huit possibilités de réponses. Suite à la création d'une version informatisée du HART, les items ont dû être adaptés à l'écran de l'ordinateur afin de permettre l'enregistrement adéquat des mouvements oculaires (Hessels *et al.*, 2011). Les modifications apportées ont eu des conséquences sur le nombre d'éléments figurant dans la matrice et sur les alternatives, qui passent à un format 2x2 avec quatre possibilités de réponse (Hessels, *et al.*, 2011).

Nonobstant ce changement, les modalités de résolution restent les mêmes. En effet, pour résoudre les items, l'enfant doit comparer les images de la première ligne, comprendre et inférer leur relation, retrouver cette même règle dans la deuxième ligne et l'appliquer ensuite à la troisième, en choisissant la bonne réponse parmi celles proposées (Hessels *et al.*, 2008). En d'autres termes, pour résoudre les items du test, l'enfant doit faire recours au raisonnement analogique.

4.2 *La phase d'entraînement*

Le HART est caractérisé par une phase d'entraînement standardisée qui enseigne à l'enfant la procédure nécessaire pour la résolution des items proposés dans le test.

Dans le paragraphe suivant, les processus principaux exercés lors de la phase d'entraînement seront présentés.

La phase d'entraînement du HART utilisée pour cette étude, inclut un travail d'apprentissage standardisé qui utilise différents exercices. Les items de la phase d'entraînement (c.f annexe E) peuvent être divisés en trois parties différentes. Dans un premier temps, l'examinateur propose

aux élèves de travailler le processus de comparaison. Il leur demande de comparer un modèle avec cinq autres figures de la même ligne. L'exercice est présenté aux élèves de manière progressive : l'examinateur cache les lignes suivantes et focalise l'attention sur une ligne à la fois. La comparaison est au centre de l'exercice : les élèves doivent confronter les différentes figures au modèle et trouver les éléments qui changent et ceux qui restent les mêmes. Cette première partie contribue aussi à travailler l'encodage et l'attention sélective : l'élève doit focaliser son attention sur les éléments qui caractérisent les deux figures ou qui les diffèrent, tout en inhibant les informations non pertinentes. L'examinateur demande donc aux élèves de relever les différences et les similitudes avant de leur donner la solution. À la fin de l'exercice, il explicite l'importance de la comparaison des figures. Dans une deuxième partie de l'entraînement, il présente une image composée de trois lignes et de huit éléments aux élèves. Premièrement, il leur demande de décrire les changements dans l'image. Il leur demande ensuite de trouver la 9^{ème} figure, tout en réfléchissant aux différents changements. Dans ce cas également, les solutions des trois exercices et leur justification sont données aux élèves après leurs réponses. Dans la dernière partie de l'entraînement, l'examinateur résout six items du test HART avec les élèves. Tout d'abord, il montre uniquement la première ligne et leur demande de décrire deux figures. Il procède de la même manière pour la deuxième ligne, mais il demande cette fois aux élèves de comparer les données de deux différentes manières : non seulement ils doivent comparer les deux figures sur la deuxième ligne, mais ils doivent aussi les mettre en relation avec les deux figures de la première. L'examinateur procède ensuite en montrant la dernière ligne. Les élèves sont encouragés à se représenter la bonne réponse sans avoir accès aux autres alternatives. Pour terminer, l'examinateur montre les réponses possibles et propose aux élèves de choisir la bonne solution en considérant la solution qu'ils avaient imaginée. Avant de justifier la bonne réponse, il leur donne deux contre-exemples avec deux fausses réponses et leur demande d'expliquer pourquoi elles ne sont pas justes.

La comparaison entre une première phase de test (le pré-test) et une deuxième phase suivant l'entraînement (le post-test) permet ainsi de mettre en évidence les possibles gains de l'apprentissage. Comme le soulignent Hessel-Schlatter et Hessel (2009a), cette phase est fondamentale pour les enfants présentant des difficultés d'apprentissage ou une déficience intellectuelle car souvent, dans les conditions d'un test classique, ils n'arrivent ni à comprendre

ce qui est attendu d'eux, ni comment ils doivent résoudre les exercices proposés. L'entraînement lors de l'évaluation du potentiel d'apprentissage est aussi très important pour les jeunes enfants et les personnes ayant une déficience intellectuelle ou une difficulté d'apprentissage, car ils n'explorent généralement pas de manière systématique dans les procédures de test et ont un faible encodage, souvent passif (Hessels & Hessels-Schlatter, 2008).

4.3 Études de validation du HART

La validité et la fidélité du HART ont été vérifiées par plusieurs études. Les résultats de deux études pilotes menées en 2002 par Berger *et al.* (2004) dans des écoles primaires, soulignent que ce test permet une meilleure évaluation de la capacité d'apprentissage des élèves par rapport à un test de QI classique et qu'il est moins influencé par des facteurs tels que le comportement, la participation et l'application en classe. En ce qui concerne la prédiction, le HART démontre un lien plus solide à la réussite scolaire qu'un test d'intelligence statique (Berger *et al.*, 2004).

Pour ce qui est des enfants « à risque », le HART permet une meilleure évaluation de leurs capacités et une distinction entre les difficultés d'ordre cognitif et les problèmes culturels et comportementaux qui ne supposent pas forcément le placement d'un enfant avec des difficultés scolaires dans une classe spécialisée (Hessels *et al.*, 2008).

Pour évaluer la validité prédictive du HART, Bider et Linder (2005) utilisent le Test dynamique de géographie (TAG) avec des enfants âgés de 6 à 12 ans, provenant d'écoles ordinaires et spécialisées. Grâce à cet outil, les élèves peuvent être évalués sur une nouvelle matière, en évitant ainsi de possibles confrontations avec des échecs liés à une branche scolaire particulière (Hessels & Tieckstra 2010). Dans ce cas, le HART montre une corrélation plus élevée avec le TAG en comparaison à un examen classique de mathématiques (Hessels & Tieckstra, 2010). En outre, les auteures mettent l'accent sur l'inadéquation des tests qui omettent une phase d'introduction pour les jeunes enfants (Bider & Linder, 2005).

Une étude similaire à celle de Bider et Linder est menée par Rumley (2007), qui utilise le Test dynamique de Chimie (TAC) avec des enfants de l'école enfantine et primaire et qui obtient, au niveau de la validité prédictive, un résultat comparable à celui de Bider et Linder. Suite à une étude concernant l'utilisation du TAC pour des adolescents avec une déficience intellectuelle légère, Hessels et Tieckstra (2010) mettent en évidence la meilleure validité prédictive du HART par rapport à celle d'un test traditionnel pour cette population. À l'aide d'une analyse de

régression, les auteurs montrent la faible corrélation du résultat du test de QI avec la réussite scolaire : le QI n'explique que 8.8% de variance dans le test de chimie. Au contraire, en ajoutant le HART à cette analyse, la prédition augmente de manière significative en atteignant 48.3 % (Hessels & Tiekstra, 2010).

Les deux études plus récentes sur le HART, celle de Rojas et Vanderlinden (2010) et celle de Schorno (2013), utilisent une version informatique du test par le biais de laquelle il est possible d'enregistrer les mouvements oculaires des participants. D'autres informations relatives au processus mis en place pendant la résolution du test peuvent ainsi être obtenues.

Dans l'étude de Rojas et Vanderlinden (2010) visant à confirmer la validité de construit du HART avec l'enregistrement des mouvements oculaires (voir ci-dessus), 11 élèves âgés de 8 à 14 ans sont soumis à une version informatique du HART qui inclut une phase d'entraînement entre le pré-test et le post-test. L'étude vise à investiguer les effets de l'entraînement sur deux groupes d'enfants : un premier comprenant des enfants ayant des difficultés d'apprentissage, et un deuxième avec des enfants sans difficultés d'apprentissage. Dans ce cas également, les résultats démontrent que l'entraînement a une influence positive sur la compréhension et la manière de résoudre les items du test. C'est-à-dire que le raisonnement analogique est vivement utilisé pour résoudre les tâches, ce qui améliore la performance des enfants de manière significative.

Schorno (2013) recourt à une version informatique du HART pour évaluer 44 élèves âgés entre 15 et 20 ans, issus d'une école genevoise spécialisée. Le but de son étude consiste en l'analyse des changements de comportement lors de la résolution du test, en analysant également les mouvements oculaires, à la suite d'un entraînement standardisé. Pour réaliser son étude, l'auteure divise l'échantillon en un groupe expérimental (GE) et un groupe de contrôle (GC). Pour le premier groupe, le HART est proposé avec la phase d'entraînement (pré-test, entraînement, post-test). Pour le deuxième, le pré-test et le post-test sont proposés sans la phase d'apprentissage. Dans ses résultats, Schorno constate une amélioration de la performance du GE dans le post-test par rapport au GC qui peut s'expliquer grâce à l'entraînement reçu.

5. Le raisonnement

5.1 *Le raisonnement inductif*

Sternberg (1986) identifie deux principaux types de raisonnement, à savoir le raisonnement inductif et le raisonnement déductif. Dans ces derniers, il est possible de distinguer trois processus : l'encodage sélectif et la comparaison sélective caractérisent le raisonnement inductif, alors que le troisième élément, la combinaison sélective, est présente dans le raisonnement du type déductif. Le raisonnement inductif est défini par Klauer (1998) comme « le processus de pensée fondamental » (p. 99). Il nous permet de découvrir une loi, une règle ou une régularité à travers l'analyse d'une situation, ce qui favorise la compréhension du monde qui nous entoure (Hessels-Schlatter & Hessels, 2009b ; Klauer, 1990).

Klauer (1990) met l'accent sur le processus de comparaison grâce auquel il est possible de détecter des propriétés communes ou des relations entre les éléments du problème, à travers des similarités et des différences. Le processus de comparaison est également fondamental pour la vérification de l'hypothèse formulée à travers une comparaison précise (Klauer, 1990).

Le raisonnement inductif montre une bonne corrélation avec les apprentissages scolaires ainsi qu'un lien avec l'intelligence générale (Klauer, 1998). Il n'est pas anodin que le raisonnement inductif ait joué, et joue encore, un rôle central dans la conceptualisation de l'intelligence ainsi que dans l'évaluation des capacités intellectuelles : il est utilisé comme mesure de l'intelligence des individus dès le début de la création des tests classiques (Goldman & Pellegrino, 1984 ; Klauer, 1990 ; Schlatter, 1999 ; Sternberg, 1977). Comme nous l'avons vu dans le chapitre précédent, le HART se base sur l'utilisation du raisonnement analogique pour résoudre les items proposés correctement. Le raisonnement analogique représente la composante la plus fréquente du raisonnement utilisé lors de la résolution de tâches inductives (Goldman & Pellegrino, 1984).

5.2 *Le raisonnement analogique*

Raisonner par analogie est une activité mentale que nous mettons en place dans des moments variés de la vie qui dépendent du contexte et de la situation dans lesquels nous nous trouvons. Ceci a lieu, par exemple, dès lors que nous nous appuyons sur une connaissance antérieure afin de comprendre ou apprendre quelque chose de nouveau (Sternberg, 1977). Le raisonnement

analogique est à la base de l'apprentissage de nouveaux concepts, de stratégies et du savoir-faire, occupant également un rôle important dans le processus de transfert des apprentissages (Schlatter, 1999). Ce raisonnement est considéré par plusieurs chercheurs comme étant une mesure significative de la pensée humaine et des capacités intellectuelles (Mulholland, Pellegrino & Glaser, 1980 ; Hessels & Hessels-Schlatter, 2010 ; Hessels & Tiekstra, 2010 ; Klauer, 1998 ; Sternberg 1977), car il est considéré comme le domaine de l'intelligence à la base du fonctionnement intellectuel, étroitement lié au *facteur g*, le facteur d'intelligence général (Berger *et al.*, 2004). Hessels-Schlatter et Hessels (2009b) définissent le raisonnement analogique comme « l'inférence d'une loi ou d'une règle à partir de l'observation de situations ou d'exemples particuliers » (p. 418).

5.3 Tâches analogiques

La structure de base d'un problème analogique à quatre termes peut être présentée de la manière suivante A : B :: C : D. Dans ce type de situation, l'individu doit être capable d'utiliser des processus spécifiques lui permettant d'obtenir la bonne réponse au problème. Les tâches analogiques se présentent sous une multitude de formats géométriques, figuratifs, numériques et verbaux. Cependant, le format n'affecte aucunement la manière de résoudre ces différentes tâches: il faut tout d'abord encoder les deux termes (A et B) en comprenant la relation qui les unit, pour ensuite la réutiliser avec le troisième terme (C), qui permettra finalement d'inférer le quatrième élément manquant (D).

Selon Schlatter (1999), deux types d'analogies sont possibles. Dans le premier cas, pour résoudre le problème, la personne s'appuie sur ses connaissances, comparables à celles du problème, de telle sorte qu'elle puisse les appliquer à cette situation. C'est donc à l'individu de choisir quels éléments seront pertinents pour la résolution et cela dépendra des expériences qu'il aura vécues. Dans le deuxième cas, la personne base son raisonnement sur les éléments donnés de la situation. Ces derniers sont nécessaires à la résolution du problème et sont présents dans l'exercice. L'individu doit donc utiliser les composantes de l'exercice. Ce deuxième type d'analogie est celui qui nous intéresse, car il est à la base de la construction d'items de plusieurs tests, aussi bien d'intelligence que d'apprentissage.

En accord avec les résultats de Bethell-Fox et de Lohman et Snow (1984), Hessels *et al.* (2011) mettent l'accent sur la relation entre la difficulté caractérisant les tâches analogiques et le nombre

d'éléments à examiner, aussi bien dans la matrice que dans les alternatives de réponses proposées. En effet, plus le nombre d'informations à retenir est élevé, plus sa mémorisation sera difficile.

5.4 Composantes en jeu

En 1904, Spearman publie un premier article dans lequel il présente les bases de sa méthode d'analyse factorielle, relative au facteur général de l'intelligence, *le facteur g*. Selon l'auteur, ce dernier serait commun à l'ensemble des tâches scolaires utilisées dans son étude (Chartier & Loarer, 2008). D'après Hessels-Schlatter et Hessels (2009b), le raisonnement ne peut être considéré comme « une entité en soi » (p. 425) et doit être conçu comme l'ensemble de plusieurs processus cognitifs. Dans sa théorie datant de 1923, Spearman présente trois processus de base impliqués dans le raisonnement inductif (Sternberg, 1986). Le premier, *apprehension of general factor*, permet d'encoder et de distinguer les informations importantes de celles qui ne le sont pas, pour la résolution de la tâche. Les éléments pertinents sont ensuite stockés dans la mémoire de travail. Le deuxième processus, *eduction of relation*, consiste en l'inférence de la relation permettant la récupération dans la mémoire à long-terme des connaissances nécessaires à la résolution de la tâche. Le troisième processus, *eduction of correlates*, consent à combiner les informations encodées et à appliquer la relation établie à un nouvel objet.

Dans son analyse, Sternberg (1978) propose la méthode « componential analysis » (p. 117) pour enquêter les processus impliqués dans l'intelligence. Pour ce faire, il étudie des tâches analogiques, qu'il sépare en quatre parties afin d'isoler les différents processus nécessaires à la résolution. Sternberg (1978) identifie cinq processus fondamentaux pour la résolution d'une analogie, chacun desquels représente une étape du traitement du problème : l'*encoding*, nécessaire pour enregistrer les caractéristiques de chaque terme de l'analogie ; l'*inference*, qui permet d'inférer la relation entre les deux premiers termes de l'analogie ; le *mapping*, qui consent à mettre en lien le premier terme avec le troisième, tout en assimilant cette relation ; l'*application*, qui consiste à utiliser la règle inférée entre les deux premiers termes pour le troisième élément, permettant ainsi de découvrir le quatrième élément ; la *response*, qui prévoit le choix de la bonne réponse parmi les alternatives proposées par l'exercice. Un sixième processus est mis en avant par Sternberg (1978) ; la *justification* de la réponse. Cette dernière est utilisée pour donner des justifications sur le choix de la réponse, dans le cas où le sujet ne serait

pas sûr de sa justesse. Goldman et Pellegrino (1984) s'intéressent également aux processus sous-jacents du raisonnement. Ces derniers, définis par les auteurs, sont pratiquement identiques à ceux de Sternberg, mis à part pour l'*inference* et le *mapping*, qu'ils considèrent comme les sous-composantes d'un seul processus : la comparaison. En outre, les auteurs ajoutent un élément de plus, celui de l'exécution, à savoir la démarche effective pour la résolution de la tâche.

Hessels-Schlatter et Hessels (2009b) présentent une liste de huit processus cognitifs nécessaires à l'utilisation correcte et efficace du raisonnement analogique, qui sont également à la base de la structuration des exercices présents dans la phase d'entraînement du HART.

Selon les auteurs, ces processus sont : *l'exploration* et *l'encodage*, nécessaires pour explorer l'exercice et enregistrer les informations ; *l'attention*, qui permet de focaliser sa concentration sur les éléments pertinents et d'inhiber ceux qui ne le sont pas; *l'abstraction de concepts*, *la comparaison*, *l'inférence* et *l'application de relations*, fondamentaux pour comprendre les informations de manière plus générale, inférer le lien entre celles-ci et construire de nouvelles relations dans les exercices. Un dernier accent est mis sur *la mémoire de travail*, processus essentiel pour la mémorisation et l'élaboration des données des exercices.

5.5 Études sur le raisonnement analogique

Dans son étude de 1989, Klauer mise à vérifier l'efficacité d'un entraînement des processus présents dans le raisonnement inductif à travers un test d'apprentissage. Les résultats sont positifs : il démontre qu'il existe un transfert des apprentissages reçus dans l'entraînement des tâches d'induction proposées. Suite à ces premiers résultats, trois programmes sont développés afin d'entraîner l'utilisation du raisonnement inductif chez les individus. Nombre d'études ont été menées sur ces programmes prenant en considération des populations de divers âges, présentant ou non des difficultés intellectuelles et ayant également un retard mental léger. Les études mettent en lumière que ce type d'entraînement donne des résultats positifs : de manière générale, les sujets entraînés augmentent leur performance ce qui n'est pas le cas des sujets non entraînés. L'apprentissage proposé dans les programmes est donc non seulement efficace mais il démontre aussi qu'il « est possible d'améliorer le raisonnement inductif de personnes présentant des troubles de l'apprentissage et un retard mental léger » (Klauer, 1998, p. 118). De plus, les effets de l'entraînement sont visibles sur le long terme, car des résultats positifs restent présents après plusieurs mois, suite à l'entraînement.

Dans sa thèse de doctorat, Schlatter (1999) propose un outil d'évaluation conçu pour les personnes ayant une déficience intellectuelle modérée à sévère : le TAPA (Test d'Apprentissage de la Pensée Analogique). Cet instrument base son évaluation sur des tâches analogiques figuratives et géométriques et vise, dans la phase d'apprentissage, à entraîner les processus cognitifs sous-jacents au raisonnement inductif. À travers ses résultats, l'auteure met en évidence la possible amélioration des capacités intellectuelles des individus testés, pouvant « faire preuve d'un certain niveau de raisonnement abstrait, contrairement à ce qui est habituellement pensé par certains chercheurs. » (Hessels-Schlatter, 2010, p. 45). Une partie des personnes participant à son étude profite de l'entraînement reçu et démontre une amélioration significative de ses capacités liées au raisonnement inductif (Hessels-Schlatter & Hessels, 2010). Cette recherche atteste une fois de plus de la possibilité d'entraîner les processus cognitifs relatifs au raisonnement inductif des personnes. Selon Hessels-Schlatter et Hessels (2009b), « [e]n agissant de manière spécifique sur les différents processus impliqués, il est possible de développer le niveau de compétence en raisonnement analogique ainsi que le fonctionnement cognitif en général » (p. 420).

6. Les mouvements oculaires

6.1 *Les yeux et leurs mouvements*

Les yeux de l'être humain recouvrent trois rôles fonctionnels (Lévy-Schoen, 1984). Le premier est un rôle considéré comme étant social : ils sont impliqués dans toutes sortes de contacts visuels lors de la communication entre deux ou plusieurs personnes. Les yeux revêtent également un rôle d'équilibre, contribuant au maintien de la stabilité de la posture du corps et à l'orientation des gestes moteurs. Le troisième rôle se situe au niveau de l'exploration de l'environnement visuel. Toutes ces actions sont rendues possibles grâce au déplacement du regard vers un objet qui nous intéresse, une action réalisée possible grâce aux mouvements oculaires.

Chaque jour, lorsque nous lisons un texte ou que nous observons une scène, nous utilisons les mouvements oculaires afin de recueillir et d'assimiler les informations du monde extérieur. Lévy-Schoen (1969) nous invite à percevoir la vision comme une *fonction active*, dans laquelle

sont impliqués différents mouvements oculaires, mis en œuvre selon le type de contexte. Gregory (1966, cité par Rayner, 1998) spécifie que les mouvements oculaires dépendent du type d'action que nous mettons en place. Par exemple, dans le cas où les yeux suivent un objet, les mouvements seront fluides et constants, alors qu'ils auront tendance à se déplacer plus rapidement lors d'une exploration visuelle.

L'observation nette du monde qui nous entoure est uniquement possible si l'image qui nous intéresse est visible par une partie spécifique de la rétine : la fovéa. Grâce aux mouvements oculaires, le regard est orienté vers un objet d'intérêt et il est maintenu avec stabilité (Zambarbieri, 2006). À l'aide de cette première description, nous pouvons identifier les deux fonctions principales des mouvements oculaires : le maintien du regard, le *gaze-holding*, et le déplacement vers une cible, la zone d'intérêt, le *gaze-shifting* (Zambarbieri, 2006). Il existe quatre principaux mouvements oculaires (Rayner, 1998). Le *pursuit* est une action qui permet à l'individu de suivre un objet mobile. Lorsque l'objet d'intérêt est extrêmement proche, il est nécessaire de focaliser son regard en bougeant les yeux l'un vers l'autre: le mouvement qui permet cette action est intitulé être *vergence*. Le *vestibular eye movement* est un mouvement oculaire qui compense le déplacement de la tête et du corps et qui consent au maintien de la vue dans la même direction. Le quatrième mouvement s'appelle *saccade*. Parmi les quatre présentés, la *saccade* est le seul mouvement qui puisse être exécuté de manière volontaire en l'absence d'un stimulus externe (Zambarbieri, 1992). Il s'agit du le mouvement oculaire le plus pertinent pour la résolution d'une tâche, et donc d'un élément fondamental à considérer dans le cadre de cette étude.

6.2 *Les fixations et les saccades*

La partie de l'œil étant capable de visualiser une image de manière optimale, se trouve sur la rétine. Bien qu'il existe deux autres zones de l'œil impliquées dans la visualisation de la cible, la zone para-fovéale et la zone périphérique, le type d'analyse mis en place par la fovéa ne peut être réalisé dans les deux autres régions (Rayner, 1998). Pour permettre à la fovéa de visualiser un objet de manière claire, les yeux doivent bouger sur la zone intéressée avec des mouvements rapides, les saccades (Van der Stigchel & Nijboer, 2011).

Entre une *saccade* et l'autre, les yeux restent pratiquement immobiles. Ce type de pause est appelé *fixation*. La durée moyenne d'une *fixation* se situe entre 40 et 500 millisecondes (Vakil,

Lifshitz, Tzuriel, Weiss, Arzuoaan, 2011). Son but premier est de maintenir la cible dans la région fovéale afin d'en faire une analyse plus fine (Rayner, 1998). Malgré ce que nous pouvons penser, les yeux ne sont jamais réellement fixés sur un objet. En effet, durant la fixation, ils mettent en œuvre trois mouvements oculaires fins: les *tremors*, les *drifts* et les *microsaccades* (Martinez-Conde, Macknik, Hubel, 2004 ; Levy-schoen, 1969).

La *saccade* est le mouvement oculaire mis en place dans les actions de la vie quotidienne, par exemple pendant la lecture, l'observation d'une scène ou la recherche d'un objet (Rayner, 1998). Le but de ce mouvement est de déplacer les yeux aussi rapidement que possible, sur un objet d'intérêt qui apparaît dans la zone périphérique de la rétine (Inchingolo, 1992 ; Yarbus, 1967). Il existe deux manières de concevoir ce mouvement: d'un côté, certains auteurs (voir par exemple Becker & Fuchs, 1969 ; Jacobson & Dowdell, 1979), le définissent comme un *ballistic movement*, un mouvement qui ne peut être corrigé une fois commencé. D'un autre côté, divers auteurs (voir par exemple Van der Stigchel, Meeter, Theeuwes, 2006) suggèrent de concevoir la *saccade* plutôt comme le vol d'un avion, avec un point de départ et un point d'arrivée précis, mais qui subit des influences pendant le voyage. Dans ce cas, la *saccade* est définie comme un « *not pre-programmed movement* » (Van der Stigchel *et al.*, 2006, p. 667), un mouvement de nature dynamique et non-linéaire. Malgré cette différence, il existe un consensus relatif au mouvement de correction mis en place après la première *saccade* : souvent, un seul mouvement n'est pas suffisant pour atteindre la cible parfaitement. Un ou plusieurs mouvements de correction sont donc nécessaires pour calibrer la cible de manière précise.

Les saccades sont réalisées par la contraction simultanée des muscles agonistes oculaires et l'inhibition des muscles antagonistes (Pettorossi, 1992). Grâce à ces deux différentes actions, les saccades peuvent atteindre une vitesse élevée, supérieure à 400°/sec et également arrêter leur trajet sur une zone d'intérêt choisie (Pettorossi, 1992). La *saccade* est le mouvement le plus rapide que le système oculaire humain puisse produire (Inchingolo, 1992) ainsi que le plus fréquent : en effet, nous produisons environ trois saccades par seconde (Zambarbieri, 2006). Ce mouvement est si rapide que, pendant cette action, l'information visuelle reçue n'est pas nette et devient donc difficile à analyser (Zambarbieri, 2006 ; Valkil *et al.*, 2010). Ce phénomène est dénommé *saccadic suppression* (Gross, Vaughan & Valenstein, 1967 ; Rayner, 1998) et décrit

un manque d'acquisition d'informations nouvelles pendant la *saccade*. Cela a pour résultat une réduction de la sensibilité visuelle et une conséquente impossibilité de percevoir une image claire. D'un point de vue quantitatif, la *saccade* est typiquement décrite à travers trois paramètres: l'amplitude, la vitesse et la durée (Becker, 1991 ; Zambarbieri, 2006). Sa vitesse et sa durée sont étroitement liées à son amplitude. En effet, si l'amplitude augmente, la vitesse et la durée le feront également (Becker, 1991). Le lien entre ces trois paramètres est dit être *main sequence* (Zambarbieri, 2006). Un quatrième paramètre évoqué par Zambarbieri (2006), la latence, représente l'intervalle de temps entre l'apparition du stimulus et le début de la saccade (Zambarbieri, 2006). Il existe par ailleurs d'autres facteurs influençant la vitesse et la durée, notamment : le mouvement directionnel de l'orbite, la manière dont la cible est désignée et le niveau attentionnel du sujet (Becker, 1991). La durée d'une saccade dépend également de la distance de la cible observée (Rayner, 1998).

6.3 Études concernant les mouvements oculaires

Dès la fin du 19^{ème} siècle, les mouvements oculaires deviennent un sujet d'intérêt parmi les scientifiques pour des raisons variées (Zambarbieri, 2006). Dans le passé, leur analyse permettait le diagnostic précoce de pathologies neurologiques (Zambarbieri, 2006) ; aujourd'hui leur utilisation est devenue un moyen de communication pour les personnes affectées par des déshabilités motrices importantes. En effet, les mouvements oculaires permettent d'écrire avec les yeux par le biais d'un clavier virtuel. Ce type d'écriture est appelé *eye typing* (Majaranta & Räihä, 2002). L'utilisation des mouvements oculaires pour la compréhension du fonctionnement des processus cognitifs est une pratique employée depuis les années 1970 (Rayner, 1998). La multiplication de ces études est rendue possible grâce aux innovations technologiques ayant permis la construction de machines et de techniques d'enregistrement moins encombrantes et plus précises. Pendant la lecture d'un texte, tout comme lors de la résolution d'une tâche analogique, le sujet met en place des stratégies variées et divers processus cognitifs. Il est très difficile de savoir exactement ce qui se passe au niveau de l'élaboration mentale, car il s'agit d'une action intellectuelle non visible à travers une simple observation du comportement. L'analyse des mouvements oculaires dans ce type de tâches est donc réellement importante, car elle permet de réfléchir de manière plus approfondie aux processus cognitifs sous-jacents et d'élargir nos connaissances relatives à la performance d'un individu (Vakil *et al.*, 2010 ;

Zambarbieri, 2006). En effet, la mesure de ces mouvements permet d'obtenir des indices précieux sur le comportement de la personne et sur les processus cognitifs qu'elle utilise pendant la résolution d'une tâche, permettant ainsi d'obtenir des informations précieuses relatives à l'analyse de la résolution du problème (Dillon, 1985 ; Lévy-Schoen, 1984).

Dans la littérature existante, de nombreuses études concernant l'écriture sont disponibles. Malheureusement, il en existe moins qui contemplent les mouvements oculaires dans la résolution de problèmes analogiques.

Concernant les études des mouvements oculaires durant la lecture, Rayner et Pollatsek (1981) démontrent que la destination de l'œil n'est pas choisie au hasard : il y a un contrôle direct sur quand et vers quoi l'œil bouge. Ce choix dépend de l'information de la *fixation* en cours et des informations de la *fixation* précédente (Rayner & Pollatsek, 1981).

Rayner (1998) souligne que la durée des fixations et la longueur des saccades dépend du type d'activité mis en place. Par exemple, dans le cas de la lecture silencieuse, la durée de la *fixation* et la longueur de la *saccade* seront moins élevées en comparaison à celles mises en place lors d'une exploration visuelle. D'autres éléments intéressants des études de Rayner sont à retenir, tel que le fait que pendant la lecture, les lettres ne sont pas toutes fixées par l'œil : il n'est donc pas nécessaire qu'elles soient toutes analysées par la zone fovéale. Rayner met également en évidence une absence de corrélation entre la longueur de la *saccade* et la durée de la *fixation* durant la lecture, alors que cette corrélation existe dans des situations qui n'impliquent pas l'élaboration linguistique de l'objet observé.

En ce qui concerne les études des mouvements oculaires en relation à la résolution de tâches analogiques, Bethge, Carlson et Wiedl (1982) utilisent le test de Raven dans le cadre d'une recherche investiguant les effets d'une procédure dynamique. Les auteurs enregistrent les mouvements oculaires de 72 enfants pendant la passation du Raven CPM, en distinguant trois groupes différents selon trois conditions de passation. Le premier groupe passe une version standard du Raven CPM. Le deuxième groupe reçoit un feedback verbal relatif à la justesse de chaque item. Le troisième groupe doit écrire et justifier, à travers la verbalisation, les éléments de la matrice et la manière lui ayant permis d'obtenir la réponse finale. Cette étude démontre l'influence d'une évaluation dynamique sur la performance au test Raven, qui augmente de

manière significative dans les conditions de feedback et de verbalisation, en réduisant également des comportements inappropriés tels que l'impulsivité et l'anxiété.

À l'aide de l'analyse des mouvements oculaires, les auteurs relèvent une différence de comportement entre un groupe et l'autre. Par exemple, dans des conditions de feedback et de verbalisation, les stratégies mises en place par les participants se démontrent être plus efficaces et davantage planifiées. En outre, dans ces mêmes conditions, les fixations augmentent aussi bien pour l'observation de la matrice que pour les alternatives.

Dans leur étude relative à la résolution de tâches analogiques géométriques, Bethell-Fox, Lohman et Snow (1984) découvrent que les stratégies utilisées par les participants lors de l'exécution sont différentes et dépendent de la complexité de l'exercice (nombre d'éléments et nombre d'alternatives présents dans l'item). À travers l'analyse des mouvements oculaires de 132 participants, les auteurs identifient deux types de stratégies mises en place pendant la résolution des tâches analogiques. La première est dite *constructive matching* et consiste en l'analyse systématique de la matrice pour se représenter la solution correcte au problème. La représentation de la solution est par la suite comparée aux alternatives proposées. Les auteurs observent que l'utilisation de cette stratégie augmente lorsque la *two-cue condition* vient à être utilisée dans la phase d'entraînement. Cette modalité consiste à montrer les deux premiers termes de l'analogie uniquement au lieu d'en montrer l'intégralité au participant. Les auteurs observent que cette première stratégie est utilisée par les participants ayant obtenu un résultat élevé. À la différence de ces derniers, les participants moins performants ont tendance à directement comparer les éléments de la matrice aux alternatives de l'exercice - surtout lorsqu'il s'agit d'items plus difficiles- en éliminant les options au fur et à mesure afin d'obtenir la bonne réponse au final. Ce type de stratégie est intitulé *response elimination*. Bethell-Fox *et al.* (1984) soulignent également l'importance de la complexité de la tâche dans le choix d'une ou de l'autre stratégie. En effet, malgré le fait que les élèves plus performants utilisent généralement la première stratégie, les moins performants ont tendance à opter pour la deuxième lorsque l'item devient plus complexe. Ce changement est dû, selon les auteurs, à l'incrémentation de la difficulté des items qui, en augmentant, entrave la capacité de la mémoire de travail dans son traitement des données.

Vigneau, Caissie et Bors (2006) utilisent les mouvements oculaires pour investiguer les raisons des différences intra et interindividuelles dans la performance de 55 étudiants universitaires pendant la passation informatisée du test Raven APM. Les principales variations entre les individus se situent au niveau de la vitesse et des stratégies utilisées pour compléter une tâche. Les auteurs soulignent une différence d'exploration de cette dernière entre les participants aux capacités faibles et ceux étant plus performants, qui se situe au niveau de l'observation de la matrice et des alternatives possibles. À l'aide des deux stratégies relevées par Bethell-Fox *et al.* (1984), les auteurs soutiennent également qu'il existe une différence dans la manière de résoudre la tâche. Les participants au score plus élevé consacrent plus de temps à l'exploration de la matrice, c'est-à-dire à l'encodage, observée de manière globale, alors qu'ils passent moins de temps à analyser les alternatives possibles. La stratégie utilisée le plus fréquemment pour l'exploration, est la *constructive matching*. Les participants ayant un score plus faible ont eux tendance à utiliser la stratégie *response elimination*, en passant ainsi moins de temps sur la matrice, qui ne vient pas à être explorée dans sa totalité. Un accent est mis sur l'influence de la difficulté dans l'exécution des tâches : plus cette dernière est présente, plus le temps général nécessaire à la résolution, l'analyse de la matrice, l'analyse des alternatives et le nombre de comparaisons, augmente.

Hessels *et al.* (2011) utilisent une version informatique du HART du type pré-test, entraînement, post-test afin d'évaluer les effets de l'entraînement sur les processus de résolution demandés par des tâches analogiques. Les mouvements oculaires de 11 enfants avec et sans difficultés d'apprentissage, ont été enregistrés. À l'aide des résultats obtenus, les auteurs soulignent un changement général de stratégie de résolution des enfants, avec ou sans difficultés d'apprentissage, suite à l'entraînement. Plusieurs différences positives entre le pré et le post-test font surface. En effet, au post-test, les enfants explorent la tâche de manière plus systématique, en consacrant plus de temps à l'analyse de la matrice et moins de temps aux alternatives de réponse. De plus, les comparaisons entre ces deux dernières se font plus rares et deviennent plus rapides et performantes. Selon les auteurs, ce résultat souligne l'importance de la stratégie *constructive matching* comme élément caractérisant une bonne performance. Cette recherche confirme l'influence de l'entraînement dans le comportement ainsi que l'importance de l'analyse des mouvements oculaires pour la compréhension des stratégies et des processus mis en place

par les participants : « [e]ye movement registration can then become a key element of educational psychology assessment, helping to identify deficient cognitive processes and strategies in children with learning difficulties » (Hessels *et al.*, 2011, p. 111).

La recherche de Schorno (2013) porte sur les changements du comportement des élèves lors de la résolution de tâches analogiques après un entraînement standardisé. Pour le groupe expérimental ayant suivi une phase d'apprentissage, la recherche montre une corrélation positive entre l'entraînement et trois variables : le score, le temps de l'encodage et le temps passé dans la matrice au post-test. Ces deux derniers indiquent un changement de comportement dans l'exécution des tâches, ce qui caractérise la stratégie de *constructive matching*. Cependant, contrairement aux résultats de Hessels *et al.* (2011), cette corrélation ne peut être retenue pour le temps total de résolution passé dans les alternatives, ni pour le nombre de comparaisons effectuées lors du post-test, car ces trois ne diminuent pas, bien au contraire. Dans certains cas ils augmentent même légèrement. Schorno fait l'hypothèse que l'entraînement amènerait les élèves, de manière générale, à consacrer plus de temps à la résolution des tâches du post-test, grâce à une analyse plus systématique des alternatives de réponse et des comparaisons entre ces dernières et la matrice. De plus, une difficulté de représentation mentale de la réponse chez les élèves pourrait expliquer ces derniers résultats. En effet, si la solution est très claire, le temps consacré aux alternatives sera amoindri. Par contre, si l'image mentale est floue, le temps passé sur les alternatives augmentera ou restera le même.

Les études présentées dans ce chapitre mettent en évidence l'importance des tests d'apprentissage pour l'évaluation intellectuelle des personnes. De plus, l'utilisation des mouvements oculaires comme des outils offrant des indices précieux sur les processus sous-jacents au raisonnement humain, est démontrée.

7. Question de recherche et hypothèses

L'objectif primaire de cette étude est d'analyser, à l'aide de l'enregistrement des mouvements oculaires, les comportements des participants pendant la résolution du test HART et de saisir d'éventuelles différences entre le pré- et le post-test. Les données seront par la suite comparés aux résultats obtenus par Schorno (2013) avec une population présentant une déficience intellectuelle. Cela nous permettra d'analyser les variations entre deux populations différentes avec le même âge mental. Grâce à ceci, une autre comparaison intéressante du comportement des personnes, à savoir avec et sans déficience intellectuelle, lors de la résolution de tâches analogiques pourra être relevée, ce qui permettra d'obtenir des informations sur la validité de construit du HART.

Les recherches précédentes ont montré que les enfants avec des difficultés d'apprentissage et les personnes présentant une déficience intellectuelle sont capables, après un entraînement, de résoudre des items du HART en mettant en œuvre les processus sous-jacents au raisonnement analogique. Pour cette recherche nous avons donc formulé les hypothèses et la question de recherche suivantes :

Hypothèses

- 1) *Les enfants et adolescents entraînés montrent de meilleures performances au post-test que les enfants et adolescent non entraînés.*
- 2) *Les enfants et adolescent entraînés montrent une meilleure utilisation des processus sous-jacents au raisonnement analogique au post-test que les enfants et adolescents non entraînés.*

Question de recherche :

Est-ce que les personnes avec une déficience intellectuelle montrent le même niveau de performance dans le HART ainsi que la même utilisation des processus de raisonnement, que des enfants et des adolescents sans difficultés, mais du même âge mental?

II PARTIE EMPIRIQUE

8. Méthodologie

8.1 Échantillon

Cette étude a été menée entre le mois de septembre et le mois de décembre 2013 dans plusieurs écoles publiques tessinoises. Les participants de cette recherche étaient au nombre de 37, dont 18 filles et 19 garçons, et vivaient dans le canton du Tessin. Leur âge chronologique était compris entre 5 ans et 10 mois et 20 ans et 5 mois, avec une moyenne de 11 ans et 2 mois. Tous les élèves avaient suivi, jusqu'au moment de la recherche, un parcours scolaire ordinaire.

Etant donné qu'il s'agissait d'une population italophone, les consignes de la phase d'entraînement du HART ont dû être traduites en italien. Afin de constituer l'échantillon, nous avons pris contact avec 12 classes de 10 institutions différentes comprenant une école enfantine, six écoles primaires, un cycle, un collège et une université.

Pour satisfaire aux exigences de cette recherche, l'échantillon a été divisé en deux groupes. Le groupe de contrôle (GC), composé de 20 élèves et le groupe expérimental (GE) dans lequel figuraient les 17 personnes restantes. Le *tableau 1* illustre l'échantillon final.

Le choix de l'échantillon n'a pas été fait au hasard. Comme la comparaison entre des personnes avec et sans déficience intellectuelle était l'un des objectifs primaires de cette étude, les personnes ont été choisies sur la base de leur âge mental et chronologique, pour que l'échantillon corresponde le plus possible à celui de Schorno (2013), tant pour le groupe de contrôle que pour le groupe expérimental.

Tableau 1. Caractéristiques de l'échantillon final

	Âge			Âge mental			Sexe	
	M	Min	Max	M	Min	Max	Homme	Femme
GE (N = 17)	11;3	7;9	20;5	11;11	7;6	20;0	8	9
GC (N = 20)	10;5	5;10	16;3	10;4	5;10	16;6	11	9
Total (N = 37)	10;10	5;10	20;5	11;2	5;10	20;5	19	18

Initialement, le nombre de participants prévus pour cette étude devait correspondre à celui de Schorno (2013), à savoir 44 élèves. Malgré nos efforts, seul 37 personnes ont pu en faire partie à final. La réduction de l'échantillon été due, pour deux personnes, à la difficulté de trouver des

élèves correspondant à ceux de Schorno avec des résultats au même niveau. Quatre personnes ont été exclues de l'étude à cause d'un enregistrement défaillant de leurs mouvements oculaires. Des soucis de santé intervenus pendant la passation du post-test du HART ont mené à l'exclusion de l'étude de la dernière personne.

Avant même de prendre contact avec les différentes écoles, un travail d'analyse de l'échantillon de Schorno (2013) a été nécessaire pour pouvoir déterminer les participants de cette recherche. En effet, afin de permettre un bon appariement de ces deux populations différentes, l'échantillon a dû répondre à plusieurs exigences : chaque personne a été sélectionnée selon un âge mental et un âge chronologique précis.

Pour déterminer l'âge mental de notre échantillon, nous avons utilisé le même instrument d'évaluation que Schorno, à savoir le test Raven SPM. L'idée était de trouver, pour chacun des 44 élèves de sa recherche, un nombre identique de personnes présentant le même âge mental.

8.2 *Instruments*

8.2.1 **Les Matrices Progressives de Raven (1998)**

Comme la majorité des tests basés sur le *facteur g*, le Raven SPM présente des exercices composés de dessins géométriques, détachés de toute signification et qui demandent l'utilisation du raisonnement analogique. L'échelle est composée de 60 exercices divisés en cinq séries (A, B, C, D et E). Chaque série contient 12 items par ordre croissant de difficulté: au fur et à mesure que l'individu avance dans une série, le nombre d'éléments compris dans la matrice, les alternatives de réponses ainsi que le degré d'analyse nécessaire à la résolution, augmentent progressivement.

Le Raven SPM, version papier-crayon, est utilisé pour la totalité des élèves dans chacune des 12 classes, permettant ensuite l'appariement des personnes selon leur âge mental.

8.2.2 **Le Hessels Analogical Reasoning Test, version informatique**

Toujours dans l'idée d'obtenir une comparaison entre deux échantillons, la version informatique du HART employée pour cette recherche est la même que celle utilisée par Schorno (2013). Il s'agit de 15 items pour le pré- et le post-test, dont dix créés par Rojas et Vanderlinden (2010) et cinq ajoutés par Schorno, issus de la banque de données du HART.

Suite à un changement de capteur des mouvements oculaires employé dans l'étude précédent la

sienne, Schorno doit apporter quelques modifications aux cinq nouveaux items pour leur utilisation dans son étude, à savoir un changement de format, permettant ainsi l'enregistrement oculaire avec le nouvel instrument. Finalement, chaque exercice présente une matrice composée de deux colonnes et de deux lignes, avec quatre alternatives de réponse. En outre, Schorno insère les chiffres 1, 2, 3 et 4 sous ces dernières, pour chaque item (c.f. Figure 1). Cela a pour but de faciliter l'enregistrement de la réponse de l'élève, qui peut utiliser les chiffres comme repères pour désigner la bonne réponse.

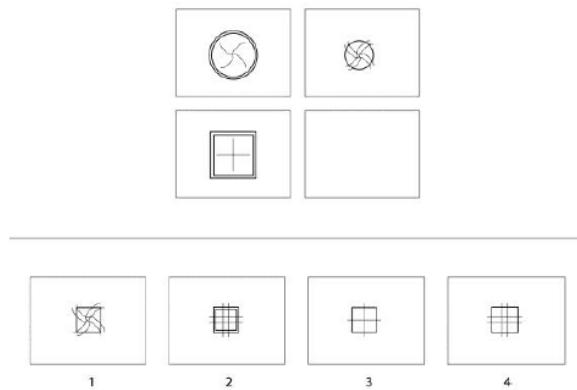


Figure 1 Item A1

Le pré-test et le post-test du HART sont deux tests parallèles comportant 15 tâches chacun. Pour chaque item du pré-test il existe un item parallèle dans le post-test, avec le même niveau de difficulté. Les paires présentent des relations et des nombres d'éléments identiques. De plus, la difficulté des items augmente progressivement dans chacun des 15 exercices (Schorno, 2013). Le HART est adapté pour son utilisation par Tobii Studio, qui permet l'enregistrement des mouvements oculaires des 37 participants, lors du pré et du post-test.

8.2.3 Eye trackers - Tobii Studio

Tobii 1750 est un dispositif d'origine suédoise permettant la mesure des mouvements oculaires. Concrètement, il s'agit d'un ordinateur portable avec un écran de 17 pouces et une résolution maximale de 1280x1024 pixels, qui contient deux caméras capables de détecter et d'enregistrer les mouvements oculaires des participants pendant la résolution du HART. Ces derniers possèdent des micro-projecteurs infra-rouges qui permettent de créer des reflets sur la cornée de l'œil, enregistrables en temps réel. L'utilisation des rayons infra-rouges rend possible un

enregistrement optimal et continu du mouvement oculaire, sans influence de la lumière externe (Tobii Technology, 2013). Les caméras, cachées par le biais d'un filtre et situées au bas de l'écran, sont pratiquement invisibles. De cette manière, la personne ne se sent pas analysée par la machine, à contrario des dispositifs *eye-tracker* mobiles, qui mesurent les mouvements oculaires grâce à un support visible et externe, comme par exemple des lunettes ou une base sur laquelle la tête de l'individu doit être posée pendant l'entier de la passation.

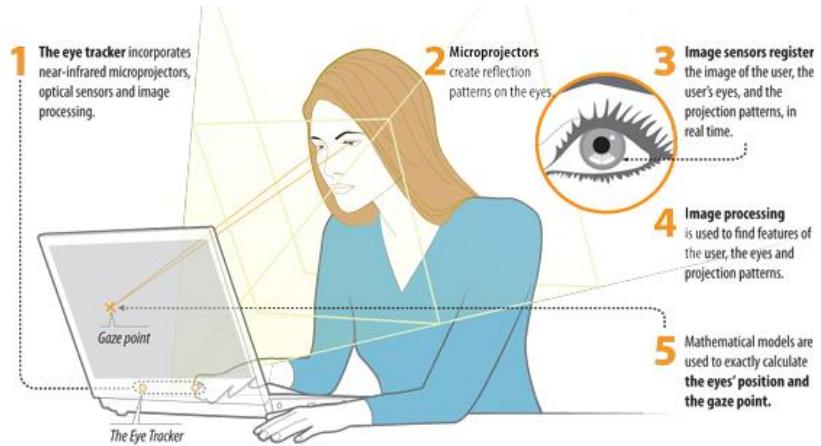


Figure 2 Système d'enregistrement des mouvements oculaires

Selon les créateurs de Tobii 1750 (Tobii Technology, 2013), l'instrument peut mesurer les mouvements oculaires dans différentes situations, comme par exemple lors d'un jeu sur l'ordinateur, d'une exploration d'un site web ou de l'analyse d'images. Tobii est capable de reconnaître les yeux, qu'il signale à l'aide d'un indicateur particulier. Par conséquent, il est également en mesure de déterminer si les yeux ne sont pas fixés sur l'ordinateur ou si la personne est mal positionnée face au dispositif (qu'elle soit trop haute, basse, proche ou éloignée). Pour cette recherche, l'outil Tobii a permis l'enregistrement des fixations des yeux et leurs changements de position pour chaque item. Le pourcentage d'enregistrement considéré par les participants se situe entre 62 et 98%. Sur 37 individus, seuls quatre d'entre eux obtiennent un enregistrement au-dessous de 70%.

Grâce à Tobii, il est possible de définir des zones particulières d'observation des mouvements oculaires lors de la réalisation des tests, ce qui est fondamental pour l'analyse ultérieure des données obtenues. Dans le cas de cette étude, ces zones, appelées Area Of Interest (AOI), sont les mêmes que celles définies par Schorno (2013). Il s'agit de huit AOI pour chaque item du HART : quatre coïncident avec les quatre cases de chaque matrice, les autres correspondent aux

quatre cases des alternatives de réponse. Plus spécifiquement, les zones relatives à la matrice sont M1 ; M2 ; M3 et M4. Leur totalité représente un premier « AOI Group » qui concerne la matrice analogique de chaque item. La même procédure a été adoptée pour nommer les quatre zones des alternatives de réponse (A1 ; A2 ; A3 et A4), qui, ensemble, forment le deuxième groupe AOI. Les deux groupes sont fondamentaux pour permettre au logiciel Tobii de pourvoir des éléments d'analyse relatifs aux mouvements oculaires (c.f. chapitre 1.4). En se basant sur une critique formulée dans des études précédentes, Schorno (2013) élargit les huit AOI de chaque item par rapport au carré encadrant chaque élément de la matrice et chaque réponse, afin de permettre une petite erreur de mesure. La phase d'entraînement n'est pas analysée dans le cadre de ce travail.

8.3 Plan de la recherche

8.3.1 Passation du Raven SPM

Dans un premier temps, une prise de contact avec les directeurs de chaque établissement a été nécessaire pour présenter la recherche et obtenir l'autorisation de travailler avec certaines classes. Suite à celle-ci, 190 élèves appartenant à 12 classes ont passé le test Raven SPM. Après cette première évaluation, 42 participants ont été sélectionnés. La différence entre les résultats des participants de cette étude et ceux de l'échantillon de Schorno consistait en un maximum de trois points pour le score brut du Raven SPM, pour l'âge mental, et à un maximum de 12 mois de différence avec l'âge chronologique par rapport à l'étalonnage déterminé par le Raven.

La définition des deux groupes, à savoir expérimental (GE) et de contrôle (GC), dépendait encore une fois de ces deux mesures. Pour chaque élève de Schorno, une personne correspondante, se rapprochant le plus possible à un participant de Schorno, selon des critères d'âge mental et chronologique, a été sélectionnée. Sur la base de ces éléments, les 42 personnes ont été divisées en deux groupes de recherche.

8.3.2 Passation du HART

Le pré-test du HART a été proposé à tous les élèves, de manière individuelle et sur l'ordinateur. En moyenne, sept jours ont passé entre le pré et le post-test, avec un minimum de cinq et un maximum de treize jours. Le pré-test a suivi les mêmes modalités de passation que le post-test.

Les deux ont été proposés à la totalité de l'échantillon, alors que la phase d'entraînement a été consacrée uniquement au GE. En d'autres mots, la modalité de passation du HART pour le GC consistait en une procédure du type pré-test / post-test, alors que pour le GE la phase d'entraînement a été rajoutée avant la passation du post-test. Le *tableau 2* illustre de manière synthétique les quatre principales phases de passation de la recherche. Dans cette deuxième étape de la recherche, le nombre des participants s'est abaissé à 37.

Tableau 2: Phases de passation de la recherche

		Phase d'entraînement (GE uniquement)
Raven SPM (classe entière)	Pré-test HART (GE et GC)	Post-test HART (GE et GC)

8.4 Procédure d'administration

8.4.1 Raven

La passation du Raven SPM a commencé à des moments différents pour chaque école, car les confirmations des institutions concernant l'autorisation de procéder à la récolte des données ne se sont pas toutes présentées en même temps. Chaque fois qu'une permission était donnée, un rendez-vous était pris avec la classe pour commencer la passation du premier test. Le Raven était toujours proposé de manière collective, à la classe entière, et aucune limitation de temps n'était prévue. Chaque élève recevait un cahier qui contenait le test, avec une feuille de réponses sur laquelle il/elle devait inscrire sa date de naissance ainsi que les réponses choisies. La présentation du test respectait la consigne standard (voir annexe A). Dans un premier temps, les élèves sont invités à ouvrir leur cahier à la première page et à suivre les instructions données par l'examinateur, pour le premier exercice. Un accent est mis sur la partie vide de la figure et sur l'objectif des exercices. Dans un deuxième temps, l'examinateur passe en revue, l'une après l'autre, les alternatives de réponse au premier exercice, en soulignant à chaque fois la raison pour laquelle elles sont plus ou moins justes. Pour finir, la réponse correcte est donnée. Avant de passer au deuxième exercice, l'examinateur explicite la manière de noter la solution sur la feuille de réponses tout en précisant que les exercices suivants se basent sur le même principe de résolution. Il souligne également que chacun peut prendre son temps pour arriver jusqu'à la fin

du cahier. Les élèves sont ensuite invités à inscrire leur réponse et à résoudre l'item suivant. Une correction de ce dernier est alors mise en place par l'examinateur, avant de laisser continuer les élèves seuls dans la résolution du test. La feuille de réponses est contrôlée pour chaque élève, afin d'éviter une confusion dans les réponses et de s'assurer que toutes les cases soient bien remplies.

8.4.2 HART

Pré-test et post-test :

Le HART a été administré de manière individuelle, dans une salle différente de celle utilisée pour les cours ordinaires. Il s'agissait de la salle de classe, de la bibliothèque scolaire ou de la classe d'un collègue absent. La salle est restée la même pour tous les élèves, aussi bien pour le pré que pour le post-test.

L'examinateur invite l'élève à s'asseoir sur une chaise face à l'instrument Tobii et l'informe sur les exercices qu'il va devoir résoudre. Avant de commencer, il demande à l'élève de veiller à ne pas trop changer de posture et d'éviter de bouger sa tête. À l'aide d'un ordinateur portable connecté à Tobii, l'examinateur calibre la position de l'élève et définit, grâce au logiciel de cet outil, si son visage est bien placé. Pour cette raison, avant chaque test, comme avant la phase d'entraînement, Tobii projette un rond rouge en mouvement sur son écran, que l'individu doit suivre avec les yeux uniquement. Le rond se déplace sur l'écran et permet au dispositif de faire la calibration des yeux de la personne. Ceci est fondamental, car si les yeux sont mal calibrés, Tobii refuse de passer à la phase du test. Une fois la calibration réalisée, chaque élève est enregistré grâce à un numéro de référence. Par ailleurs, la requête présentée dans les consignes de passation (c.f annexe B), concernant la limitation des mouvements du corps, a dû être justifiée afin de mettre les personnes plus à l'aise face à ce nouvel instrument. L'examinateur explique donc à l'élève que, selon des recherches, les personnes parviennent à mieux résoudre les exercices si elles maintiennent une bonne posture. C'est pour cela, qu'une fois assis correctement, l'élève est invité à limiter ses mouvements. Si, pendant la résolution, le calibreur des mouvements oculaires signale un mauvais enregistrement causé par un changement de posture de la part de l'élève, l'examinateur pourra l'inviter à se rasseoir correctement, en le renvoyant à sa posture initiale. L'introduction de cette justification auprès de l'élève a été très efficace, surtout pour les enfants plus jeunes, car cela leur a donné une « excuse » au fait de devoir rester dans la même position

pendant toute la phase du test.

Avant de commencer la passation, l'examinateur fournit une explication relative aux items et à la manière dont l'élève devra donner sa réponse. Entre chaque item du pré et du post-test, un point de fixation, sous forme de rond gris, apparaît au centre de l'écran et reste visible pendant deux secondes, pour ensuite passer à l'item suivant. L'examinateur prévient l'élève qu'il y aura un point et lui demande de le fixer jusqu'à ce qu'il ne disparaisse pas. Le passage entre le point de fixation et l'item suivant est automatique, ce qui facilite l'enregistrement des mouvements oculaires, car l'enregistrement de chaque item peut toujours commencer au même endroit, au centre de l'écran.

L'enregistrement des mouvements oculaires débute tout de suite, dès l'item d'introduction du test, grâce auquel le participant a la possibilité de se familiariser avec les exercices du HART avant de commencer le test. À ce moment, l'examinateur décrit la matrice à travers le nombre de colonnes, de lignes et de figures présentes (c.f. annexe B). Comme pour le Raven SPM, il est expliqué à l'élève qu'il existe qu'une seule réponse correcte pour chaque exercice. Après la description de l'item, l'examinateur invite l'élève à proposer la réponse qu'il considère être juste. Ensuite, indépendamment de la justesse de sa réponse, l'examinateur passe en revue les quatre alternatives, en expliquant les raisons pour lesquelles elles sont correctes ou non.

Suite à ce premier exemple, la véritable phase de test peut commencer. Chaque item est projeté sur l'écran de Tobii et l'élève dispose de tout le temps nécessaire pour donner sa réponse. L'examinateur note celle-ci sur une feuille, en évitant ainsi au participant de devoir déplacer son regard de l'écran. Afin de passer d'un item à l'autre, il est nécessaire d'appuyer sur une touche du clavier ou de cliquer à l'aide de la souris. Ce geste est réservé à l'examinateur, afin d'éviter une réaction impulsive de la part de l'élève. En effet, une fois que l'item suivant est lancé, il n'y a plus de possibilité de revenir en arrière. Après l'énonciation verbale de la réponse choisie, l'élève donne la commande à l'examinateur de changer d'exercice. Le participant doit résoudre les 15 items du test (voir annexe F) sans l'aide de l'examinateur, qui reste néanmoins présent pour noter les réponses et pour lancer l'item suivant.

Phase d'entraînement

La phase d'apprentissage, uniquement réalisée pour le groupe expérimental, avant son passage

au post-test, vise à enseigner à l'élève la manière de résoudre des problèmes impliquant le raisonnement analogique. En premier lieu, l'examinateur présente différents exercices permettant d'entraîner le processus de comparaison, l'inférence, l'attention sélective et la représentation mentale de l'élève (voir annexe E). Cette série d'exercices est ensuite suivie de six items à résoudre, pratiquement identiques à ceux présents dans le HART. En ce qui concerne la phase d'entraînement, également construite sur l'ordinateur, compatible et projetable sur l'écran de Tobii, son enregistrement n'a pas été considéré dans les analyses. En effet, dans cette partie du test, l'examinateur interagit avec l'élève et influence ses mouvements oculaires, car il guide son attention vers les images à l'écran. Un enregistrement valide des mouvements oculaires était alors presque impossible et peu intéressant pour les finalités de cette étude. Une description plus détaillée de la phase d'entraînement est présentée dans le chapitre 4.2 de ce travail.

8.5 Méthode d'analyse des données

L'analyse des résultats de cette étude a été possible en considérant principalement trois types de données. D'une part, nous avons retenu les scores des différents tests, classés selon le nombre de réponses correctes. D'autre part, nous avons analysé les mouvements oculaires à l'aide de plusieurs critères découlant directement du logiciel Tobii Studio. Pour finir, nous avons observé les différences de performance aux tests, ainsi que les mouvements oculaires impliqués dans la résolution des tâches du HART, entre les enfants et les adolescents sans difficultés intellectuelles et ceux ayant une déficience intellectuelle, en utilisant l'échantillon de Schorno (2013).

8.5.1 Les variables relatives aux scores des différents tests

Afin d'utiliser les données des trois tests utilisés, nous nous sommes basés sur les résultats finaux de chaque participant de l'échantillon. Le Raven SPM et le HART utilisent la même modalité de cotation des réponses : si la personne choisit la bonne solution, un point lui est assigné. Au contraire, si la réponse donnée est fausse, aucun point n'était attribué. Le test Raven donne un maximum de 60 réponses justes, alors que le total des exercices du HART correspond à 15 pour le pré-test et 15 pour le post-test.

8.5.2 Tobii studio

Comme expliqué dans le chapitre 8.2, la réalisation informatique du HART avec le logiciel Tobii a impliqué la construction de zones d'intérêt (AOI). Ces huit zones, prises individuellement ou

dans un AOI group, ont permis l'enregistrement et l'observation de plusieurs variables relatives aux mouvements oculaires aussi bien dans le pré-test que dans le post-test. Il s'agissait de six éléments qui donnaient des informations différentes sur la manière dont les participants avaient analysé les informations à l'écran. Cela nous donnait par conséquent un indice intéressant sur les comportements mis en place, nous permettant de nous renseigner sur les différences entre les deux groupes de l'échantillon. Parmi ces six variables, cinq provenaient de l'outil *Statistic - Visit duration*, lequel consentait à directement obtenir des données traitées statistiquement par le logiciel. La dernière variable, également issue du logiciel Tobii studio, par le biais du programme *Visualisations - Gaze Plot*, a nécessité un travail d'analyse supplémentaire pour pouvoir recevoir des informations intéressantes dans le cadre de cette étude. Ci-dessous, chaque variable de Tobii Studio sera présentée de manière plus complète.

Visit duration

Premier temps d'encodage de la matrice (Encodage). Ce premier élément correspond au temps passé avant la première fixation d'une AOI dans les alternatives. Schorno (2013) a mis en évidence la limite de cette mesure, car elle n'enregistre pas le réel temps d'encodage, à savoir le temps que l'individu consacre uniquement aux éléments de la matrice avant de déplacer son regard sur la première AOI des alternatives de réponse, mais calcule plutôt la tranche de temps qui s'écoule du début de l'enregistrement de l'item jusqu'au moment où les yeux passent aux alternatives. Cet outil mesure bien évidemment le temps d'encodage, mais inclut également le temps pendant lequel l'individu ne regarde pas les zones d'intérêt de la matrice.

Temps passé sur la matrice (Matrice). Les quatre zones d'intérêt relatives à la matrice sont impliquées ici. Cette variable calcule le temps total que les yeux de chaque participant passent sur la matrice de chaque item pour toute la durée de l'enregistrement.

Temps ailleurs (Ailleurs). Il y a des moments lors desquels les participants observent l'item sans se focaliser sur les huit AOI. Le temps passé ailleurs calcule cela. Pour ce faire, Tobii enregistre le temps que chaque individu passe en dehors des huit zones d'intérêt.

Temps sur les alternatives (Alternatives). Les AOI utilisées pour l'enregistrement de cette quatrième variable, consistent en quatre zones relatives aux alternatives de réponse (A1 ; A2 ; A3 ; A4), permettant ainsi de calculer le temps total consacré à l'analyse des alternatives de chaque item pendant le test.

Temps total (temps total). Cette dernière variable se distingue des quatre premières pour deux raisons. Premièrement, sa mesure se fait en secondes (s), alors que les autres sont mesurées en millisecondes (ms). Deuxièmement, cette variable ne peut être directement enregistrée par le logiciel Tobii, car elle représente l'ensemble de plusieurs enregistrements, à savoir le temps consacré aux alternatives additionné à celui passé sur les alternatives et le temps passé ailleurs. Comme Schorno (2013) le soulève dans sa recherche, Tobii Studio ne permet pas de calculer le temps total du test, mais plutôt le temps total de l'enregistrement oculaire. Cette mesure ne tient donc pas compte du temps lors duquel le regard des participants n'est pas capté par les caméras.

Gaze Plot

Ce deuxième outil a permis d'obtenir une image des items du pré- et du post-test, indiquant le parcours des yeux, du début à la fin de chaque exercice. Les saccades sont illustrées par une ligne, les fixations sont définies à travers des ronds de différentes dimensions : plus la durée est longue plus le rond s'agrandit. Plusieurs images sont présentées dans l'annexe G. À l'aide de cet outil, il a été possible de compter le nombre de comparaisons entre la matrice et les alternatives, réalisées par chaque participant. L'outil Gaze Plot n'en offrait cependant pas le chiffre total, mais indiquait simplement le trajet exécuté. C'est pourquoi le calcul des comparaisons a été effectué manuellement, en comptant le nombre de passages réalisés par les yeux entre les éléments de la matrice et les alternatives dans chacun des 30 items du HART. Il nous semblait important de souligner une difficulté de précision dans le calcul de ces trajets effectués par les yeux, pour certains items. En effet, Gaze Plot indiquait la totalité des trajets par item, dont le nombre total devenait difficile à saisir dans le cas où le nombre de comparaisons était élevé, car souvent plusieurs trajets étaient superposés (c.f. annexe G). Dans certains cas, nous avons donc fait recours à l'enregistrement vidéo des mouvements oculaires, ce qui nous a permis d'observer et de compter les trajets plus clairement, car seul un trajet à la fois était visible. Le nombre de comparaisons correspond à la variable *Comparaisons MA* dans les tableaux d'analyses.

Comparaison de deux échantillons

Afin de saisir les différences de performance et de comportement entre les personnes avec et sans déficience intellectuelle lors de la résolution de tâches analogiques, nous avons mené des analyses multi-variées.

Les deux groupes ont été considérés comme quatre sous-groupes, à savoir le GE (GEd) et le GC (GCd) de Schorno (2013), ayant une déficience intellectuelle et le GE (GEo) et le GC (GCo) de base, « construit » pour cette étude. Nous nous sommes basés sur les résultats finaux obtenus par chaque participant de l'échantillon.

9. Présentation des résultats

La présentation des résultats se divise en deux parties principales. Il y aura en premier lieu l'analyse des scores aux tests pour l'échantillon de base, composé du GCo et du GEo. Elle sera suivie de la présentation des résultats relatifs aux variables de temps et de comportement, provenant de l'enregistrement des mouvements oculaires. La deuxième partie sera consacrée à la comparaison des résultats des deux échantillons, à savoir ceux des participants avec et sans déficience intellectuelle, sur ces mêmes variables.

9.1 Analyse de l'échantillon de base

9.1.1 Scores aux tests

Dans cette première partie de l'analyse, les scores aux différents tests utilisés au cours de la recherche seront pris en considération afin de saisir les changements de performance des 37 participants composant l'échantillon final. Le but principal est celui de voir les effets de l'entraînement sur les résultats des participants, pour déterminer s'il existe ou non un gain au niveau du post-test pour le GEo, dû à l'entraînement. Le *tableau 3* illustre la moyenne obtenue au pré- et au post-test du HART pour les deux groupes de recherche, le résultat de l'analyse de variance ainsi que la taille d'effet (η_p^2) et la puissance. La taille d'effet est considérée comme étant petite pour des valeurs allant de .01 à environ .05, les valeurs à partir de .06 caractérisant une taille d'effet moyenne et les valeurs égales à ou plus élevées que .13 étant considérées comme grandes (Cohen, 1992). En vue de comprendre le rôle joué par la variable *groupe de recherche* sur la performance des élèves, une analyse statistique de covariance est réalisée. Le résultat au post-test du HART est considéré comme la variable dépendante, celui du pré-test est compris comme covariable, et la variable *groupe*, (GEo vs. GCo), comme facteur fixe.

Tableau 3. Moyennes, Ecarts-types, F, df, p, aux pré- et post-tests du HART des 37 élèves

Score au test	Prétest HART				Post-test HART				<i>F</i>	<i>df</i>	<i>p</i>	η_p^2	Puissance						
	GCo N = 20		GEo N = 17		GCo N = 20		GEo N = 17												
	M	SD	M	SD	M	SD	M	SD											
	7.2	2.2	7.9	2.2	8.5	2.8	10.2	2.3	3.421	1	.07	.09	.44						

Une différence de moyenne dans le score brut du Raven SPM existe entre les deux groupes. Sur

un total de 60 points, le GEo obtient une moyenne de 40.2 (SD = 8.3). Le GCo, quant à lui, comptabilise 36.6 points avec un écart-type de 8.5. Cette différence n'est pas significative et a une taille d'effet petite ($F_{3,77} = 1.01$, ns , $\eta_p^2 = .04$, power = .26)

Concernant le score au test HART illustré dans le *tableau 3*, une légère différence de 0.7 points existe pour le GEo au pré-test. Cette disparité augmente d'un point (1.7) exactement au post-test : le GEo obtient un score moyen de 10.2 points, alors que celui du GCo est de 8.5 points. L'analyse de variance (ANOVA) de ces résultats nous indique que l'effet de groupe est marginalement significatif, que la puissance est faible, avec une taille d'effet moyenne de la variable *groupe*, correspondant à .09.

Dans le *tableau 4*, la corrélation de Pearson est utilisée pour investiguer le lien possible entre la performance au test Raven SPM et celle au HART.

Tableau 4. Corrélation de Pearson entre Raven SPM, âge chronologique, Hart pré et post-test pour GC et GE

	AC	HART pré-test	HART post-test
SPM GCo	.90**	.74**	.82**
SPM GEo	.93**	.68**	.58*

* $p \leq 0.5$ ** $p \leq 0.1$

La variable *âge chronologique* (AC) est également ajoutée à cette analyse. Pour le GCo, la corrélation est significative et forte pour les trois variables ($p \leq .001$), positive et haute pour l'âge chronologique, pour le pré et le post-test. Pour le GEo également, le Raven SPM corrèle toujours avec le pré-test, le post-test et l'âge chronologique, de manière positive. Cependant, un léger changement est à mettre en évidence au niveau de la signification et de la puissance. En effet, pour le post-test, la corrélation avec le SPM reste toujours positive et significative ($p \leq .05$), mais la force diminue, passant de forte à moyenne. Cela signifie que l'entraînement a changé l'ordre des participants du GEo au niveau du score.

9.1.2 Analyse des données de Tobii Studio

Dans le *tableau 5*, les scores du pré- et du post-test du HART sont analysés avec les résultats relatifs aux variables de temps, à savoir le premier temps d'encodage (*Encodage*), le temps consacré à l'analyse de la matrice (*Matrice*), celui passé dans les alternatives de réponse

(*Alternatives*), le temps passé ailleurs (*Ailleurs*) et le temps total (*Temps total*), composé par la somme de *Matrice*, *Ailleurs* et *Alternatives*. La variable dépendante est toujours le score de la variable au post-test, avec la variable *groupe* comme facteur fixe. La variable de temps au pré-test est toujours considérée comme covariable. Le nombre de comparaisons entre la matrice et les alternatives (*Comp.-MA*) est également analysé.

Sur la base de ces données, il est possible de déterminer d'éventuels changements des mouvements oculaires chez les participants de l'étude, suite à l'entraînement.

Encodage

Lors du pré-test, le GCo passe moins de temps que le GEo à analyser les éléments de la matrice avant de passer aux alternatives. La différence entre les deux groupes est de 12s. Suite au post-test, le temps d'encodage augmente pour les deux groupes, à savoir d'environ 9s pour le GCo et de 17s pour le GEo. La différence n'est pas significative, la taille d'effet est petite.

Matrice

Une différence de temps passé sur la matrice est déjà présente au pré-test. Elle est plus élevée de 6s pour le GCo. Au post-test, les deux groupes montrent une augmentation de temps par rapport à la moyenne obtenue au pré-test. Celle-ci est plus haute pour le groupe ayant suivi l'entraînement (respectivement 2.2s pour le GCo et 3.3s pour le GEo). L'analyse de variance de ces résultats nous indique que l'effet de groupe n'est pas significatif et que la taille d'effet est petite.

Alternatives

Pour ce qui concerne le temps passé sur les alternatives au pré-test, les deux groupes y consacrent en moyenne un temps similaire, soit respectivement 12.6s pour le GCo et 12.4s pour le GEo. Au post-test, le temps augmente pour les deux groupes, passant à une moyenne de 13.5s pour le GCo et de 13.9s pour le GEo. La différence n'est pas significative et la taille d'effet, comme la puissance, sont à nouveau petites.

Ailleurs

La variable temps ailleurs est également non significative, avec une taille d'effet très petite et une faible puissance. Cette variable diminue légèrement au post-test (0.2s) pour le GCo, alors qu'elle augmente légèrement pour le GEo (0.5s).

Tableau 5. Moyennes et Écarts-types de l'échantillon final pour le HART et données statistiques pour chacune des variables

	Pré-test HART				Post-test HART				<i>F</i>	<i>df1</i>	<i>df2</i>	<i>p</i>	η_p^2	Puissance							
	GCo N = 20		GEO N = 17		GCo N = 20		GEO N = 17														
	M	SD	M	SD	M	SD	M	SD													
<i>Latences*</i>																					
Encodage	38.785	17.422	51.141	19.688	47.353	21.485	68.361	37.103	.759	1	34	.39	.02	.14							
Matrice	18.173	4.572	21.481	4.910	20.362	5.909	24.805	7.598	.579	1	34	.45	.02	.12							
Alternatives	12.638	2.263	12.421	3.127	13.489	2.064	13.860	2.675	1.246	1	34	.27	.04	.19							
Ailleurs	3.988	.940	3.208	.758	3.752	.792	3.662	3.273	.312	1	34	.58	.01	.8							
Temps total	34.792	4.146	37.110	7.055	37.602	5.522	42.764	10.202	2.212	1	34	.15	.06	.30							
<i>Latences proportionnelles</i>																					
Matrice	52.4	8.1	56.9	6.4	53.9	8.7	56.7	8.6	.042	1	34	.82	.00	.06							
Alternatives	37.3	7.8	33.1	5.3	37.2	8.2	33.3	5.5	.184	1	34	.67	.01	.07							
Ailleurs	10.4	2.2	10.7	2.3	8.9	1.9	10.0	6.7	.606	1	34	.44	.02	.12							
<i>Comportement</i>																					
Comp.-MA	121.800	22.759	139.647	34.975	116.316	33.358	149.177	43.724	3.182	1	34	.084	.09	.41							

*en secondes (s) sauf Comp.-MA, nombre de comparaisons

Temps total

Le Temps total augmente pour les deux groupes au post-test, mais celui du GCo est deux fois plus élevé que celui du GCo (2.8s contre 5.7s). Ces différences ne sont pas significatives, la taille d'effet est moyenne, la puissance reste faible.

Comparaisons entre matrice et alternatives

Déjà lors du pré-test, la moyenne du nombre de comparaisons entre les analogies de la matrice et les alternatives de réponse est légèrement plus élevée pour le GCo. Au post-test, les deux groupes obtiennent des résultats inverses : les comparaisons augmentent pour le GCo et sont au nombre de 10, alors que le GCo diminue les siennes de 6. La différence n'est pas significative, la taille d'effet est moyenne, la puissance observée est basse.

Malgré le fait que les différences ne soient pas significatives, en général, la moyenne des indices de temps et de comportement tend à augmenter davantage pour le GCo que pour le GCo.

Les tableaux suivants illustrent les corrélations unilatérales entre les scores au pré- et au post-test du HART. Le *tableau 6* se réfère au GCo, alors que *le tableau 7* illustre les corrélations du GCo. Le but de ces résultats est d'indiquer s'il existe un lien entre les variables de Tobii Studio et la performance au test HART.

Tableau 6:Corrélations de Pearson (signification unilatérale) entre pré- et post-test avec données Tobii pour GCo

	Latences					Comportement	
	Encodage	Matrice	Alternatives	Temps total	Ailleurs	Comp. -MA	
Pré-test HART	.30	.42*	-.44*	.08	-.70**	.34	
Post-test HART	.23	.21	-.30	.09	-.53*	.35	

*p ≤ 0.5 **p ≤ 0.1

Pour le pré-test du GCo, trois corrélations ont une force modérée, deux sont positives (*Matrice* et *Comp.-MA*) et une est négative (*Alternatives*). Le premier temps d'encodage montre une corrélation faible et positive. La variable *Temps total* est positive alors que celle intitulée *Ailleurs* est négative, les deux faisant preuve d'un lien fort. Trois variables montrent une corrélation

significative avec le pré-test du HART, à savoir *Matrice*, *Alternatives* ($p < .05$) et *Ailleurs* ($p < .001$). Au post-test, les variables *Encodage* et *Temps total* maintiennent le même pattern, la même force ainsi que la même direction qu'au pré-test. La variable *Matrice* n'est plus significative et diminue en force, en passant à une puissance faible, en maintenant toujours la même direction. La variable *Alternatives* change également son pattern ($p = 1.0$), tandis que sa force et sa direction restent les mêmes qu'au pré-test. Une seule variable montre un lien significatif avec le score du post-test, à savoir le temps passé ailleurs ($p < .05$). Comme dans le pré-test, la corrélation est également négative et sa force s'amoindrit, en passant de forte à moyenne. Le nombre de corrélations entre la matrice et les alternatives reste faible et positif, mais sa signification devient marginale ($p = .07$).

Pour le pré-test du GGeo, les variables *Temps total* et le nombre de comparaisons (*Comp.-MA*) révèlent une corrélation négative très faible. Le temps passé sur la matrice a une corrélation positive et faible. Les trois variables analysées ne sont pas significatives. Comme pour le GCo, les variables *Alternatives* et *Ailleurs* sont significatives (respectivement, $p = .01$ et $p = .03$) et corrèlent négativement avec le pré-test, ayant une force moyenne. La seule variable désignant un lien fort, positif et significatif lors du pré-test est *Encodage* ($p < .05$).

Tableau 7:Corrélations de Pearson (signification unilatérale) entre pré- et post-test avec données Tobii pour GGeo

	Latences					Comportement
	Encodage	Matrice	Alternatives	Temps total	Ailleurs	
Pré-test HART	.62*	.28	-.51*	-.08	-.48*	-.10
Post-test HART	.50*	.17	-.03	.08	-.12	10

* $p \leq 0.5$ ** $p \leq 0.1$

Au post-test, les variables *Matrice* et *Encodage* maintiennent leur pattern, ainsi que leur direction. *Matrice* conserve également sa force (faible), alors que celle d'*Encodage* diminue, en passant de forte à modérée. Les variables *Temps total* et *Comparaisons* changent de direction en allant de positive à négative, et sont toujours non significatives avec une force faible. En ce qui concerne les variables *Alternatives* et *Ailleurs*, leur corrélation avec le post-test du HART n'est plus significative. De plus, la force du lien passe de modérée à très faible.

Le premier temps d'encodage est la seule variable qui corrèle significativement avec le score du post-test. ($p = .02$).

9.2 Analyse des deux échantillons, avec et sans déficience intellectuelle

9.2.1 Scores aux tests

Concernant les scores au Raven SPM, il n'y a pas de différences significatives entre l'échantillon avec déficience intellectuelle et celui sans déficience intellectuelle. Les scores du HART, des deux échantillons, sont analysés grâce à une analyse de variance de mesures répétées. Les résultats au pré- et au post-test représentent le facteur temps, les groupes comme facteurs fixes, ainsi que l'interaction entre groupe et performance, c'est-à-dire l'évolution des scores entre le pré- et le post-test.

Au niveau des moyennes, la performance de tous les groupes augmente au post-test. Le test multi-varié signale un effet de temps (Wilks' $\lambda = .63$, $F_{1,77} = 45.600$, $p < .001$, $\eta_p^2 = .37$, power = 1.00), mais pas un effet d'interaction entre temps et groupe (Wilks' $\lambda = .94$, $F_{3,77} = 1.675$, ns, $\eta_p^2 = .06$, power = .42). Les groupes sont significativement différents : $F_{3,77} = 3.868$, $p < .05$, $\eta_p^2 = .13$, power = .81). En effet, les tests post-hoc mettent en évidence que le GCd est marginalement différent du GEd ($p = .058$) et que ce dernier est différent du GCo ($p < .05$). Le graphique de la *figure 3* illustre l'évolution des quatre groupes du pré au post-test. Le GCd et le GCo réalisent une performance plus faible au pré et au post-test. Les groupes de contrôle évoluent ainsi dans la même direction. La meilleure performance au pré- et au post-test est celle du GEd, suivie de celle du GCo. Pareillement aux groupes de contrôle, les deux groupes expérimentaux évoluent de la même manière.

9.2.2 Analyse des données de Tobii Studio

Comme pour l'analyse des scores du HART, une analyse de variance de mesures répétées est employée. Les variables *Encodage*, *Matrices*, *Alternatives*, *Ailleurs*, *Temps* et *Comp.-MA* au pré- et au post-test constituent le facteur temps, les groupes comme facteur fixe et l'évolution des groupes comme interaction.

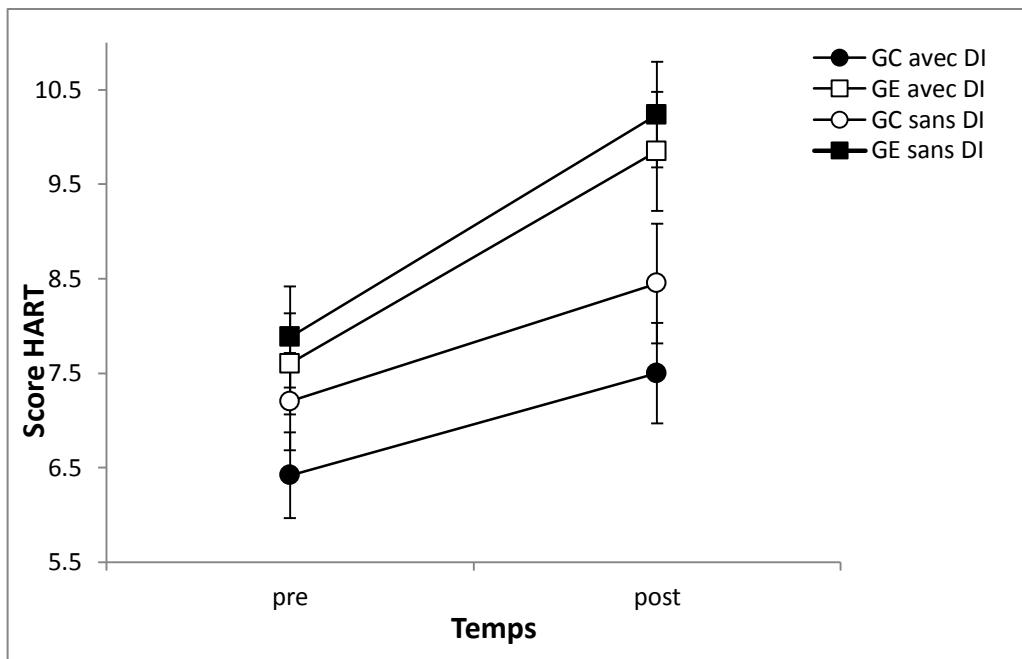


Figure 3 Évolution du score au HART pour GCd, GEd GCo, GEo.

Encodage

L’analyse multi-variée affiche un effet de temps significatif (Wilks’ $\lambda = .80$, $F_{1,77} = 19.530$, $p < .001$, $\eta_p^2 = .20$, power = 0.99) ainsi qu’une interaction significative entre le temps du pré- et celui du post-test et le groupe (Wilks’ $\lambda = .84$, $F_{3,77} = 4.895$, $p < .01$, $\eta_p^2 = .16$, power = .90). Quant au facteur groupe, il n’est pas significatif et la taille d’effet est moyenne ($F_{3,77} = 1.767$, ns, $\eta_p^2 = .06$, power = .44). En général, le GCd voit son premier temps d’encodage diminuer au post-test, alors que celui des trois autres groupes (respectivement GEd, GCo, GEo) augmente, le GEo obtenant la durée la plus élevée.

Alternatives

Concernant le temps moyen passé sur les alternatives de réponse, ni l’effet de temps (Wilks’ $\lambda = .98$, $F_{1,77} = 1.484$, ns, $\eta_p^2 = .02$, power), ni l’effet d’interaction (Wilks’ $\lambda = .95$, $F_{3,77} = 1.503$, ns, $\eta_p^2 = .06$, power = .38) ne sont significatifs. Les groupes sont significativement différents ($F_{3,77} = 10.469$, $p < .001$, $\eta_p^2 = .29$, power = 1.00). L’échantillon sans déficience passe amplement moins de temps sur les alternatives que celui avec déficience intellectuelle (respectivement GCd et GEd vs. GCo : $p < .01$; GCd et GEd vs. GEo, $p < .001$). En général, les groupes maintiennent le même niveau du pré- au post-test et suivent une évolution presque identique. Le GEd passe cependant plus de temps sur l’analyse des alternatives dans le post-test, ce qui est aussi le cas du

GEo et du GCo, qui restent néanmoins plus rapides que le groupe avec déficience. Contrairement aux autres groupes, qui augmentent le temps passé sur les alternatives au post-test, le GCd est le seul qui devient plus rapide.

Matrice

Le temps passé sur la matrice n'indique pas de différences significatives (Wilks' $\lambda = .97$, $F_{1,77} = 2.084$, ns , $\eta_p^2 = .03$, power = .30). L'effet d'interaction est à la limite de la signification et la taille d'effet se situe entre moyenne et grande (Wilks' $\lambda = .91$, $F_{3,77} = 2.684$, $p = .052$, $\eta_p^2 = .10$, power = .63). La différence de temps des quatre groupes est significative et la taille d'effet est grande ($F_{3,77} = 4.727$, $p < .01$, $\eta_p^2 = .16$, power = 0.88). Les tests post-hoc montrent que le GGeo est significativement différent du GEd ($p < .05$) et GCd ($p < .01$). En général, dans l'échantillon avec déficience intellectuelle, le GC uniquement, diminue le temps passé sur la matrice. Pour l'échantillon sans déficience intellectuelle, le GC et le GE augmentent tous deux, le GE prenant plus de temps pour l'analyse de la matrice.

Ailleurs

L'effet de temps passé ailleurs n'est pas significatif, avec une taille d'effet grande (Wilks' $\lambda = .99$, $F_{1,77} = 1.116$, ns , $\eta_p^2 = .02$, power = .18). Concernant l'effet d'interaction, les groupes se développent de la même manière dans le temps (Wilks' $\lambda = .97$, $F_{3,77} = .708$, ns , $\eta_p^2 = .05$, power = .19). L'effet de groupe dans le temps montre une différence significative, avec une taille d'effet grande ($F_{3,77} = 42.643$, $p < .001$, $\eta_p^2 = .62$, power = 1.0). Les tests post-hoc indiquent une différence significative entre les deux échantillons (GGeo et GCo versus GEd: $p < .001$; GGeo et GCo versus GCd: $p < .001$). En général, pour l'échantillon avec déficience intellectuelle, les deux groupes diminuent légèrement le temps passé ailleurs que dans les zones d'intérêt. Sur les quatre groupes, le GEd est celui qui passe le plus de temps ailleurs. L'échantillon sans déficience intellectuelle, présente un niveau beaucoup plus bas dès le pré-test. Au fur et à mesure, le GGeo diminue légèrement son temps ailleurs, alors que le GOc l'augmente légèrement.

Temps total

Pour cette dernière variable, l'effet de temps montre une différence non significative, avec une taille d'effet très petite (Wilks' $\lambda = .99$, $F_{1,77} = .606$, ns , $\eta_p^2 = .01$, power = .12), comme c'est le cas de l'effet d'interaction qui indique par contre une taille d'effet moyenne (Wilks' $\lambda = .93$, $F_{3,77}$

$= 2.049$, ns , $\eta_p^2 = .07$, power = .51). L'effet de groupe est significatif ($F_{3,77} = 14.978$, $p < .001$, $\eta_p^2 = .37$, power = 1.0). Les tests post-hoc montrent une différence significative entre l'échantillon avec et celui sans déficience intellectuelle. Précisément, le GEd est significativement différent du GGeo et du GCo ($p < .001$), comme cela l'est pour le GCd (respectivement $p < .05$ pour GGeo et $p < .001$ pour GCo). En général, les participants du GEd consacrent plus de temps à l'exécution du post-test HART. Le GCd a un niveau de départ similaire à celui du GEd, légèrement plus haut au pré-test, mais diminuant au post-test, tout en restant considérablement haut. Pour l'échantillon sans déficience intellectuelle, le temps général est beaucoup moins élevé que pour l'échantillon avec déficience intellectuelle. Les deux groupes suivent une évolution similaire, c'est-à-dire que le temps total augmente légèrement au post-test.

Comparaisons entre matrice et alternatives

La comparaison entre la matrice et les alternatives illustre un effet de temps non significatif, avec une taille d'effet petite (Wilks' $\lambda = .96$, $F_{1,65} = 3.064$, ns , $\eta_p^2 = .05$, power = .41), et un effet d'interaction non significatif (Wilks' $\lambda = .92$, $F_{3,65} = 1.862$, ns , $\eta_p^2 = .08$, power = .46). Il existe une différence significative dans les groupes, avec une taille d'effet grande ($F_{3,65} = 5.149$, $p < .05$, $\eta_p^2 = .19$, power = .91). Les tests post-hoc montrent une différence significative ($p < .05$) entre les moyennes du GEd et du GGeo. En général, Le GEd réalise beaucoup plus de comparaisons du pré- au post-test. Le GCd et le GGeo augmentent également le nombre de comparaisons au post-test, mais l'évolution du pré- au post-test reste moins forte. Quant au GCo, il s'agit du groupe qui se livre le moins aux comparaisons dès le pré-test, ainsi que du seul qui diminue légèrement leur nombre au post-test.

III DISCUSSION ET CONCLUSION

10. Discussion

Le but de cette recherche était d'investiguer les effets de la phase d'entraînement d'un test d'apprentissage sur deux échantillons différents : un premier composé d'enfants et d'adolescents, un deuxième formé par des adolescents avec le même âge mental que les premiers et présentant une déficience intellectuelle (Schorno, 2013). Deux objectifs principaux ont motivé cette étude. D'une part, nous nous sommes penchés sur l'observation des changements de performance ainsi que sur l'utilisation des mouvements oculaires pour le premier échantillon. D'autre part, nous avons analysé la comparaison de ces résultats avec le deuxième échantillon, afin d'investiguer sa manière de performer et de raisonner.

La question de recherche soulevée portait sur l'observation des différences possibles entre les deux échantillons au niveau de la performance au HART, ainsi qu'à celui de l'utilisation de processus sous-jacents au raisonnement, visibles à l'aide des mouvements oculaires.

Dans les paragraphes qui suivent, nous discuterons des deux hypothèses qui ont guidé cette étude. Nous essaierons ensuite de répondre à notre question de recherche.

À l'aide des données relatives à la performance, il est possible de constater que les élèves de l'échantillon de base ne réalisent pas de haute performance au pré-test, avec un score qui correspond en moyenne à 7.55 points sur 15, un résultat équivalent à environ la moitié du score maximal. La moyenne de l'échantillon de base augmente à 9.35 points au post-test, ce qui signifie que l'échantillon exécute en moyenne 62% des exercices correctement, que nous pouvons comparer aux 50% du pré-test. Comme pour le cas des élèves présentant une déficience intellectuelle, illustré par plusieurs recherches, (p. ex : Schorno, 2013 ; Hessels, 2009, Klauer, 1998), ce résultat montre non seulement qu'une procédure de type classique ne résulte pas nécessairement en une bonne performance générale des élèves, mais aussi que l'introduction d'une phase d'entraînement standardisée permet une augmentation du score et une meilleure performance. Concernant les deux groupes de l'échantillon de base, les résultats montrent que les enfants et adolescents sans déficience intellectuelle, ayant suivi une phase d'entraînement, obtiennent de meilleurs scores au post-test que les enfants et adolescents non entraînés. Les élèves du GEo obtiennent un score de 1.7 points de plus que ceux du GCo, au post-test. Il faut également souligner une amélioration générale des deux groupes au post-test, qui pourrait s'expliquer par un effet de répétition. Si nous observons l'effet de l'entraînement du HART pour

le GÉo à travers l'analyse de covariance, nous pouvons déduire qu'il est marginalement significatif ($p = .07$). Pour cette raison, la première hypothèse peut être partiellement confirmée. Une autre donnée intéressante, liée à la performance au HART, concerne la corrélation de Pearson. En effet, à la différence du post-test du GCo, qui montre une corrélation positive et haute avec le Raven SPM, la corrélation du GÉo est moins forte et moins significative. Cela indique que, plus le score au SPM est haut, plus les élèves du GCo sont performants au post-test du HART, alors que ce lien est moins fort et significatif pour les élèves du GÉo. Ces résultats sont tout à fait attendus et positifs pour cette recherche. En effet, dans le cas du GCo, une haute corrélation entre le Raven SPM et le HART est souhaitée, car il s'agit de deux tests extrêmement similaires qui devraient donner les mêmes résultats. La diminution de la corrélation entre le SPM et le post-test pour le GÉo, est également un résultat positif : il s'agit d'un test d'apprentissage qui, comme prévu, explique le changement de l'ordre des participants au niveau du score, car une phase d'entraînement est insérée.

La deuxième hypothèse formulée porte sur le recours de l'échantillon de base, aux processus sous-jacents du raisonnement analogique. Selon notre postulat initial, au post-test, les enfants et adolescents entraînés devraient montrer une meilleure utilisation des mouvements oculaires, que les enfants et adolescents non entraînés, surtout lorsqu'il s'agit d'actions impliquées dans la stratégie de type *constructive matching* (Bethell-Fox *et al.*, 1984 ; Mulholland *et al.*, 1980, Hessels *et al.*, 2011), comme par exemple le premier temps d'encodage. Cette deuxième hypothèse ne peut être confirmée. En effet, malgré le fait que toutes les variables indiquent une différence de moyenne entre les deux groupes, ce qui va dans le même sens que les résultats de Schorno (2013), le groupe de recherche ne semble pas influencer significativement ces variables. Ces résultats peuvent s'expliquer par une importante variance entre les scores des deux groupes. En outre, s'agissant d'un échantillon passablement petit, les gains ne peuvent être considérés comme étant significatifs. De plus, suite à une imprécision lors de l'enregistrement des mouvements oculaires de quelques participants, certaines données de la variable *Ailleurs* ne peuvent être utilisées (sur un total de 1'110 enregistrements, 28 items sont inutilisables). Le calcul de la moyenne des participants qui n'ont pas cette variable dans tous les items est alors remplacé par celui de la moyenne totale des items mesurés, en évitant ainsi de fortement biaiser le résultat.

Cependant, une analyse de corrélation avec signification unilatérale nous permet de déduire une influence importante de certaines variables sur l'échantillon. Au pré-test, le GCo montre un lien significatif pour les variables *Matrice*, *Alternatives* et *Ailleurs*. Ces deux dernières, auxquelles s'ajoute la variable *Encodage*, sont également significatives au pré-test du GEo ($r = -.51$; $r = -.48$). Au post-test, seul le temps passé ailleurs reste significatif pour le GCo, alors que l'unique variable avec un lien significatif et modéré pour le GEo, est le temps de premier encodage. Le lien entre le temps passé sur les alternatives et le temps ailleurs disparaît suite au post-test pour le GEo, en devenant non significatif, avec une taille d'effet très petite ($r = -.03$; $r = -.12$). Le pattern de corrélation entre la variable du premier temps d'encodage et le post-test du HART témoigne de l'existence de changements significatifs pour le GEo, qui pourraient être dus à l'entraînement.

Cette deuxième partie de la discussion porte sur la question de recherche, relative à la comparaison des résultats des deux échantillons, contenant des personnes avec et sans déficience intellectuelle.

En ce qui concerne la performance, l'analyse de variance multi-variée permet de relever un développement similaire pour les groupes expérimentaux et les groupes de contrôle du pré- au post-test : leur performance augmente, celle des deux groupes expérimentaux se situant plutôt au haut de l'échelle, et celle des deux groupes de contrôle, étant plus basse. Les quatre groupes s'améliorent, leurs niveaux se rapprochent (respectivement pour les deux GE et pour les deux GC), indépendamment de la présence d'une déficience intellectuelle (c.f. *figure 3*). Malgré la taille d'effet moyenne, au vu de la présence de l'effet d'interaction non significatif, nous ne pouvons associer cette amélioration à l'impact de l'entraînement. Au final, le GEo obtient la meilleure performance, suivi du GEd. Les groupes de contrôle obtiennent quant à eux un résultat plus bas.

Les variables obtenues grâce à l'enregistrement des mouvements oculaires nous permettent d'en venir à la conclusion que le temps sur les alternatives, le temps passé ailleurs, le temps total, tout comme le nombre de comparaisons entre la matrice et les alternatives, ne montrent pas un effet de temps ni un effet d'interaction significatifs pour le groupe de recherche. Cela signifie que les différences de variables entre le pré- et le post-test ne sont pas imputables à l'entraînement. En ce qui concerne la variable du temps passé sur la matrice, il existe une différence de moyenne entre les groupes au post-test, qui augmente de manière générale. La signification marginale de

l'effet d'interaction suggère que cette différence ne peut être attribuée à l'entraînement avec certitude (Wilks' $\lambda = .91$, $F_{3,77} = 2.684$, $p = .052$, $\eta_p^2 = .10$, power = .63). La seule variable qui présente une augmentation de temps au post-test ainsi qu'un effet d'interaction significatif (Wilks' $\lambda = .84$, $F_{3,77} = 4.895$, $p < .01$, $\eta_p^2 = .16$, power = .90) est celle du premier temps d'encodage. Cela signifie que les groupes emploient un temps qui varie d'un groupe à l'autre, lors du premier encodage de la matrice. Ils ne se développent donc pas de la même manière. À l'exception du GCd, le temps d'encodage des groupes augmente au post-test, surtout pour le GEd, qui est suivi du GEd et du GCo. Ce changement peut être attribué à l'influence de l'entraînement sur le temps d'encodage des groupes, élément fondamental pour l'application de la stratégie *constructive matching* (Bethell-Fox *et al.*, 1984). Comme Schorno (2013), ainsi que d'autres auteurs (Mulholland *et al.*, 1980) le relèvent avant notre étude, ce résultat pourrait s'expliquer par un manque d'automatisation de la stratégie apprise lors de l'entraînement, stratégie menant les individus à être plus performants, mais aussi moins rapides à cause d'une possible surcharge de la mémoire de travail lors de la mise en place du *constructive matching*.

Discussion générale

De manière générale, les principaux résultats de cette étude démontrent que, suite à l'entraînement, les élèves sans déficience intellectuelle obtiennent une meilleure performance au HART que les élèves sans déficience qui sont non entraînés. Une amélioration est néanmoins également visible chez les personnes présentant une déficience intellectuelle. La signification marginale présente dans les analyses de l'échantillon de base peut s'expliquer par une variance élevée, due aux participants ayant obtenu des résultats extrêmes, qui biaisent ainsi la variance. En outre, l'analyse multi-variée des deux échantillons suggère qu'il n'existe pas de différences au niveau de l'évolution des groupes pour ce qui est de la performance. Malgré le fait que les scores finaux du post-test soient légèrement moins élevés pour l'échantillon avec une déficience intellectuelle, les deux groupes expérimentaux évoluent de la même manière, à savoir en améliorant leur résultat. Cette observation est également vraie pour les groupes de contrôle : leur évolution dans le temps est extrêmement similaire, même si l'augmentation du score demeure moins importante que celle des élèves entraînés. Ce résultat peut être comparé à celui de Schorno (2013), ainsi qu'à ceux d'autres études (voir par exemple : Bider & Linder, 2005 ; Hessels, 2009 ; Rojas & Vanderlinden, 2010 ; Rumley, 2007).

En ce qui concerne les processus impliqués dans le raisonnement analogique pour l'échantillon de base, visibles par le biais de l'enregistrement des mouvements oculaires, les effets de l'entraînement ne donnent pas lieu à des résultats significatifs pour ces variables au post-test. Pourtant, la corrélation entre le premier temps d'encodage et le post-test du GÉo nous permet de déduire l'importance de cette variable sur la performance de l'échantillon de base. Contrairement à ce que Schorno (2013) expose dans son étude, la variable *Encodage* incarne, au final, le seul comportement significatif qui se modifie suite au post-test et qui est imputable à l'impact de l'entraînement.

Pour ce qui est de l'observation des mouvements oculaires des deux échantillons, le premier temps d'encodage uniquement, change significativement au post-test, ce qui est attribuable à l'entraînement. Dans le cas du temps total passé sur la matrice, il est également possible d'établir une interaction entre cette variable et l'entraînement reçu. Cependant, cette relation est à la limite de la signification. Quant à *Encodage* et *Matrice*, leur temps augmente au post-test pour les groupes expérimentaux. Ces deux variables sont les plus importantes lors de la résolution de tâches analogiques, car elles jouent un rôle central dans l'application de la stratégie dite de *constructive matching*, employée pour son efficacité et la performance vers laquelle elle mène (Bethell-Fox *et al.*, 1984). Cette augmentation peut s'expliquer par une analyse plus systématique et précise de la part des élèves ayant suivi l'entraînement et qui, pour ce faire, doivent mettre en œuvre une analyse nouvelle, non automatisée et donc, plus lente (Schorno, 2013). Toutefois, les autres variables prises en compte, à savoir le temps dans les alternatives, le temps total et le nombre de comparaisons entre matrice et alternatives, ne semblent pas être influencées par l'entraînement.

11. Limites et perspectives nouvelles

11.1 Limites

Deux limites relatives à la phase de récolte des données sont ressorties lors de cette recherche.

La modalité informatisée, d'enregistrement des mouvements oculaires, consistait en une méthode simple et pratique d'obtenir des informations riches et fondamentales pour la compréhension de la manière dont les personnes réfléchissent et mettent en place les processus impliqués lors de la

résolution. Malgré cela, pour certains individus, la récolte de ces informations s'est montrée compliquée. La première limite s'est traduite par une difficulté dans la récolte de données d'enfants très jeunes, à savoir de l'école enfantine et des premières années de l'école primaire. En effet, ces élèves n'étaient pas habitués à rester immobiles pendant longtemps, face à un écran, en évitant de trop déplacer leur regard et en devant produire un travail individuel, sur l'ordinateur. Certains enfants étaient en effet mal à l'aise avec la modalité de l'exercice et intimidés par cette situation extraordinaire qui les obligeait à travailler face à une personne inconnue. De plus, pour les plus jeunes s'est ajoutée la difficulté de donner la bonne réponse à l'aide de quatre chiffres, car ils n'avaient pas encore assimilé les nombres. Étant donné qu'ils n'étaient pas sûrs de l'appellation de ces derniers, ils répondaient en montrant du doigt la réponse choisie, ce qui influençait la qualité des mouvements oculaires. La modalité d'enregistrement des mouvements oculaires s'est donc révélée ne pas être efficace pour cette population.

Une autre limite a été la taille réduite de l'échantillon. L'objectif initial était celui d'obtenir le même nombre de participants que pour l'étude de Schorno (2013), mais cela n'a pas été réalisable à cause de plusieurs obstacles explicités dans le chapitre 8.1. Plus l'échantillon est de petite taille, plus il est difficile d'obtenir des résultats significatifs, car les participants aux résultats extrêmes ont plus d'influence sur la variance.

11.2 Perspectives

En obtenant des résultats marginalement significatifs, il serait intéressant d'améliorer cette étude en élargissant l'échantillon de recherche de manière générale et en prenant en considération essentiellement des données relatives aux jeunes enfants. Cela permettrait d'avoir moins de variance dans les résultats et de pouvoir ainsi confirmer l'influence de certaines variables qui se sont uniquement montrées être marginalement significatives. Une autre amélioration relative à la variance consisterait à tenter de combler les enregistrements des mouvements oculaires manquants avec des données heuristiques qui en permettraient l'interpolation.

Il serait également enrichissant d'élaborer une meilleure manière d'enregistrer les mouvements oculaires des enfants les plus jeunes. Une possibilité serait de leur faire porter des lunettes ou de réduire les mouvements de leur tête grâce à un dispositif fixe (Zambarbieri, 2006). Cependant, nous considérons que cette solution ne serait pas tout à fait adéquate, car cette procédure

s'éloignerait aussi bien des modalités de passation des autres élèves que de la situation d'apprentissage ordinaire.

En général, cette étude a permis d'entrer en contact avec différentes écoles tessinoises et de pouvoir discuter avec les enseignant(e)s et les professionnels du soutien pédagogique. Toutes ces personnes ont montré de l'intérêt pour ce type d'évaluation, ainsi que pour l'instrument utilisé. Afin que les tests d'apprentissage puissent trouver une place importante dans les pratiques scolaires, nous considérons fondamental que l'interaction et l'échange avec les écoles soit également maintenu dans le cadre des prochaines études sur ce sujet.

12. Conclusion

Cette recherche, dont l'objectif est de mieux comprendre les capacités intellectuelles et la manière de procéder face à des tâches analogiques pour des personnes avec ou sans déficience intellectuelle, s'inscrit dans la lignée des travaux souhaitant mettre en évidence l'importance d'une évaluation dynamique (comme par exemple Berger *et al.*, 2004 ; Bethge *et al.*, 1982 ; Bider & Linder, 2005 ; Hessels, 2009 ; Hessels *et al.*, 2011; Rumley, 2007 ; Schlatter, 1999).

Le test d'apprentissage HART a été utilisé dans une procédure d'évaluation dynamique à l'aide de l'enregistrement des mouvements oculaires, pour comparer deux échantillons distincts, à savoir un formé par des adolescents présentant une déficience intellectuelle et un autre constitué d'enfants et d'adolescents sans déficience intellectuelle. La comparaison de la performance de ces deux échantillons nous a permis de souligner que les enfants et adolescents sans déficience intellectuelle et entraînés, profitait de l'entraînement de la même manière que les adolescents présentant une déficience intellectuelle. Ce constat va dans le même sens que les études existantes (voir par exemple Rojas & Vanderlinden, 2010 ; Klauer, 1998), qui ont souligné l'importance de la stimulation et de l'entraînement du raisonnement logique chez les personnes présentant une difficulté intellectuelle ou un trouble de l'apprentissage, mais aussi chez les enfants sans difficultés. L'observation des mouvements oculaires des participants a mis en évidence le rôle fondamental de la variable *Encodage* pour la résolution ainsi que pour la performance. Les données relatives aux autres variables testées se lisaien moins clairement et se

sont montrées difficilement interprétables. Une amélioration de cette étude, à savoir l’élargissement de l’échantillon constitué, pourrait certainement avoir pour résultat des données intéressantes pour ces variables relatives au comportement.

Cette étude a révélé qu’une procédure de type classique n’offrait pas nécessairement une estimation précise de la performance réelle des personnes en matière de raisonnement analogique. L’insertion d’une phase d’entraînement, caractérisant les procédures d’évaluation dynamique, soulignait donc une fois de plus la nécessité de cette étape pour assurer une évaluation correcte des capacités de raisonnement de toute personne, indépendamment d’un déficit intellectuel ou de difficultés scolaires. De plus, grâce à l’aide des mouvements oculaires, nous avons pu mettre en évidence l’importance d’une évaluation plus complète, pouvant nous renseigner sur les comportements mis en œuvre pendant la résolution.

Cette étude a permis, une nouvelle fois, d’insister sur l’importance des tests d’apprentissage, non seulement pour l’obtention d’une meilleure évaluation des capacités intellectuelles, mais aussi pour l’observation, à l’aide de l’analyse des mouvements oculaires, de la manière de procéder et de résoudre des tâches. Cet aspect est fondamental, aussi bien pour une évaluation standardisée, comme celle visée par les test d’apprentissage, que pour un travail clinique, car dans les deux cas, il permet un cadre plus précis des processus de pensée et une intervention ciblée et spécifique aux besoins de chacun.

Comme illustré dans nombre de recherches, le raisonnement analogique peut être entraîné et modifié suite à un entraînement spécifique. Notre recherche, malgré ses limites, a permis de l’attester une fois de plus.

IV REFERENCES BIBLIOGRAPHIQUES

1. Bibliographie

- Becker, W. (1991). Saccades. In R.H.S. Carpenter (Ed.), *Eye movements* (Vol. 8, pp.95-137). London: The Macmillan Press.
- Becker, W. & Fuchs, A. F. (1969). Further Properties of the human saccadic system: eye movements and correction saccades with and without visual fixation point. *Visual Research*, 10, 1247-1258.
- Beckmann, J. F. (2006). Superiority: Always and everywhere? On some misconceptions in the validation of dynamic testing. *Educational & Child Psychology*, 23, 35-49.
- Berger, J.-L., Bosson, M., & Hessels, M.G.P. (2004). Evaluer le potentiel d'apprentissage. Construction d'un instrument à passation collective. *Pédagogie spécialisée* (2), 21-27.
- Bernaud, J.-L. (2009). *Test et théories de l'intelligence*. Paris: Editions DUNOD.
- Bethell-Fox, C.E., Lohman, D.F., & Snow, R.E. (1984). Adaptive reasoning: Componential and eye movement analysis of geometric analogy performance. *Intelligence*, 8(3), 205-238.
- Bethge, H-J., Carlson, J., & Wiedl, K.H. (1982). The effects of dynamic assessment procedures on Raven Matrices performance, visual search behavior, test anxiety and test orientation. *Intelligence*, 6, 89-97.
- Bider, C., & Linder, A.-L. (2005). *Contribution à la validation d'un test d'apprentissage, le HART (Hessels Analogical Reasoning Test), à l'aide de critères dynamiques*. Mémoire de licence [non publié] soutenu à l'Université de Genève, Faculté de psychologie et des Sciences de l'Education.
- Bosson, M. (2008). *Acquisition et transfert de stratégies au sein d'une intervention métacognitive pour des élèves présentant des difficultés d'apprentissage*. Thèse de doctorat en Sciences de l'Education, Université de Genève.
- Büchel, F. (2007). *L'intervention cognitive en éducation spéciale. Deux programmes métacognitifs*. Université de Genève : Carnets des sciences de l'éducation.
- Budoff, M. (1967). Learning potential among institutionalized young adult retardates. *American Journal of Mental Deficiency*, 72, 404-411.

- Budoff, M. (1987a). The Validity of Learning Potential Assessment. In C. Lidz (Ed.), *Dynamic Assessment: An Interactional Approach to Evaluating Learning Potential* (pp. 52-81). New York: Guilford Press.
- Budoff, M. (1987b). Measures for Assessing Learning Potential. In C. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 173-195). New York: Guilford Press.
- Budoff, M., & Friedman, M. (1964). « Learning potential » as an assessment approach to the adolescent mentally retarded. *Journal of consulting psychology*, 28(5), 434-439.
- Campione, J. C., & Brown, A. L. (1987). Linking Dynamic Assessment with School Achievement. In C. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 82-115). New York: Guilford Press.
- Carlson, J.S., & Wiedl, K.H. (1995). Principes de l'évaluation dynamique: l'application d'un modèle spécifique (F.P. Büchel, trad.). In F.P. Büchel (Dir.), *L'éducation cognitive. Le développement de la capacité d'apprentissage et son évaluation*. Neuchâtel, Suisse : Delachaux & Niestlé. (Original publié 1992).
- Chartier, P., & Loarer, E. (2008). *Evaluer l'intelligence logique. Approche cognitive et dynamique*. Les outils du psychologue, Dunod : Paris.
- Cohen, J. (1992). Statistics a power primer. *Psychology Bulletin*, 112, 155–159.
- Dias, B. (2001). Evaluation du potentiel d'apprentissage. In Doudin, P.-A., Martin, D. & Albanese, O. *Métacognition et éducation, aspects transversaux et disciplinaires*. (pp. 130 - 143). Bern: Peter Lang SA.
- Dillon, R.F. (1985). Eye Movement Analysis of Information Processing under Different Testing Conditions. *Contemporary educational psychology*, 10, 387-395.
- Feuerstein, R., Rand, Y., & Hoffman, M. (1979). *The dynamic assessment of retarded performers: The learning potential assessment device (LPAD)*. Baltimore: University Park Press.
- Flammer, A., & Schmid, H. (1995). Tests d'apprentissage: concepts, réalisation, évaluation (F.P. Büchel, trad.). In F.P. Büchel (Dir.), *L'éducation cognitive. Le développement de la capacité*

d'apprentissage et son évaluation. Neuchâtel, Suisse : Delachaux & Niestlé. (Original publié 1982).

Goldman, S.R., & Pellegrino, J.W. (1984). Deductions about induction: Analyses of Developmental and Individual Differences. In R.J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 2, pp. 149-197). Hillsdale, NJ: Erlbaum.

Gross, E. G., Vaughan, H. G. & Valenstein, E. (1967). Inhibition of visual evoked responses to patterned stimuli during voluntary eye movements. *Electroencephalography and Clinical Neurophysiology*, 22, 204-209.

Mulholland, T.M., Pellegrino, J.W., & Glaser, R. (1980). Components of geometric analogy solution. *Cognitive Psychology*, 12(2), 252-284.

Haywood, H. C (1977). A cognitive approach to the education of retarded children. *Peabody Journal of Education*, 54, 110-116.

Hessels, M.G.P. (1997). Low IQ but High Learning Potential: Why Zeyneb and Moussa do not belong in Special Education. *Educational and Child Psychology*, 14, 121-136.

Hessels, M.G.P. (2009). Estimation of the predictive validity of the HART by means of a dynamic test of geography. *Journal of Cognitive Education and Psychology*, 8(1), 5-21.

Hessels, M.G.P., Berger, J.-L., & Bosson, M. (2008). Group assessment of learning potential of pupils in mainstream primary education and special education classes. *Journal of Cognitive Education and Psychology*, 7, 43-69.

Hessels, M.G.P. & Hessels-Schlatter, C. (2008). Pedagogical Principles Favouring the Development of Reasoning in People with Severe Learning Difficulties. *Educational and Child Psychology*, 25, 66-73.

Hessels, M.G.P., & Hessels-Schlatter, C. (2010). L'apport des tests d'apprentissage dans l'évaluation des capacités intellectuelles des élèves avec difficultés scolaires. In M.G.P. Hessels & C. Hessels-Schlatter (Eds.), *Evaluation et intervention auprès d'élèves en difficultés* (pp. 5 – 34). Berne, Suisse : Peter Lang.

Hessels, M.G.P., Vanderlinden, K., & Rojas, H. (2011). Training effects in dynamic assessment: A pilot study of eye movement as indicator of problem solving behavior before and after

- training. *Educational & child psychology*, 28(2), 101-113.
- Hessels, M.G.P. & Tieksstra, M. (2010). Evaluation des capacités de raisonnement et prédition des apprentissages dans un domaine scolaire nouveau chez des élèves présentant une déficience intellectuelle légère. In Hessels, M. G. P. & Hessels-Schlatter, C. *Evaluation et intervention auprès d'élèves en difficultés*. (pp. 51-66). Bern : Peter Lang SA.
- Hessels-Schlatter, C. (2010). Les jeux comme outils d'intervention métacognitive. In Hessels, M. G. P. & Hessels-Schlatter, C. *Evaluation et intervention auprès d'élèves en difficultés*. (pp. 99–128). Bern : Peter Lang SA
- Hessels-Schlatter, C., & Hessels, M.G.P. (2009a). Clarifying some issues in dynamic assessment: Comments on Karpov and Tzuriel. *Journal Of Cognitive Education And Psychology*, 8(3), 246-251.
- Hessels-Schlatter, C. & Hessels, M. G. P. (2009b). Principes pédagogiques favorisant le développement du raisonnement. In Guerdan, V., Petitpierre, G., Moulin, J.-P., & Haelewyck, M.-C. *Participation et responsabilités sociales. Un nouveau paradigme pour l'inclusion des personnes avec une déficience intellectuelle*. (pp. 417-427). Bern : Peter Lang.
- Huteau, M. & Lautrey, J. (1999). *Evaluer l'intelligence*. Paris : Puf.
- Huteau, M. & Lautrey, J. (2006). *Les tests d'intelligence*. Paris : La Découverte
- Inchingolo, P. (1992). Modelli matematici del meccanismo saccadico. In Traccis, S. et Zambarbieri, D. *I movimenti saccadici*. (pp. 125-183). Bologna: Patròn Editore.
- Klauer, K.J. (1989). Teaching for analogical transfert as a means of improving problem-solving, thinking and learning. *Instructional Science*, 18, 179-192.
- Klauer, K.J. (1990). A process theory of inductive reasoning tested by teaching of domain-specific thinking strategies. In F.P. Büchel & J.-L. Paour (Eds.), Assessment of learning and development potential: theory and practices (Special issue). *European Journal of Psychology of Education*, 2, 191-205. 74.
- Klauer, K.J. (1998). Entraîner le raisonnement inductif chez les enfants en difficultés d'apprentissage et ayant un retard mental léger. In F.P. Büchel, J.-L. Paour, Y. Courbois, & U. Scharnhorst (Eds.), *Attention, mémoire, apprentissage. Etudes sur le retard mental*

- (pp.97-118). Lucerne, Suisse : Editions SZH/SPC.
- Lévy-Schoen, A. (1969). *L'étude des mouvements oculaires. Revue des techniques et des connaissances*. Paris: Dunod.
- Lévy-Schoen, A. (1981). Flexible and/or rigid control oculomotor scanning behavior. In D.F. Fischer, R.A. Monty, & J.W. Senders (Eds.), *Eye movement: Cognition and visual perception* (pp. 299-314). Hillsdale, NJ: Erlbaum.
- Majoranta, P., Räihä, K-L. (2002). Twenty Years of Eye Typing: Systems and Design Issues. *Journal of eye movement Research*, 2, 15 – 22.
- Martinez-Conde, S., Macknik S. L., Hubel, D. H. (2004). *The Role of Fixational Eye Movements in Visual Perception*. *Nature Reviews, Neuroscience*, 5, 229-240.
- Paour, J-L., Cèbe, S. (1999). Le mouvement de l'éducation cognitive. In Doudin, P.-A., Martin, D. & Albanese, O. *Métacognition et éducation, aspects transversaux et disciplinaires*. (pp. 107 - 140). Bern : Peter Lang SA.
- Pettorossi, V., E. (1992). Fisiologia dei movimenti saccadici: controllo sottocorticale. In Traccis, S. et Zambarbieri, D. *I movimenti saccadici*. (pp. 25-37). Bologna: Patròn Editore.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124 (3), 372-422.
- Rayner, K. & Pollatsek, A. (1981). Eye movement control during reading : evidence for direct control. *The Quarterly Journal of Experimental Psychology*, 33A, 351-373.
- Rey, A. (1934). D'un procédé pour évaluer l'éducabilité : Quelques application en psychopathologie. *Archives de Psychologie*, 24, 297-337.
- Rojas, H., & Vanderlinden, K. (2010). *Effets d'un entraînement du raisonnement analogique: Analyse du mouvement oculaire d'enfants avec et sans difficultés d'apprentissage*. Mémoire de licence [non publié] soutenu à l'Université de Genève, Faculté de Psychologie et des Sciences de l'Education.
- Rumley, L. (2007). *Etude de la validité prédictive d'un test d'apprentissage : le HART*. Mémoire de licence [non publié] soutenu à l'Université de Genève, Faculté de psychologie et des Sciences de l'Education.

Schlatter, C. (1999). *Le «test d'apprentissage de la pensée analogique» (TAPA): fondements théoriques et empiriques*. Thèse de doctorat en Sciences de l'Education, Université de Genève.

Schorno, S. (2013). *Analyse de l'effet de l'entraînement du raisonnement analogique : utilisation du capteur du mouvement oculaire dans une procédure d'évaluation dynamique*. Mémoire de licence soutenu à l'Université de Genève, Faculté de psychologie et des Sciences de l'Education.

Sternberg, R.J. (1977). Component processes in analogical reasoning. *Psychological Review*, 84(4), 353-378.

Sternberg, R.J. (1986). Toward a unified theory of human reasoning. *Intelligence*, 10, 281-314.

Tobii Technology. (2013). *Learn more about eye tracking*. Retrieved 14 may, 2014, from <http://www.tobii.com/en/about/what-is-eye-tracking/>

Vakil, E., Lifshitz, H., Tzuriel, D., Weiss, I., & Arzuoa, Y. (2011). Analogies solving by individuals with and without intellectual disability : Different cognitive patterns as indicated by eye movements. *Research in Developmental Disabilities*, 21, 846-856.

Van der Stigchel, S., Meeter, M., & Theeuwes, J. (2006). Eye movement trajectories and what they tell us. *Neuroscience and Biobehavioral Reviews*, 30, 666-679.

Van der Stigchel, S. & Nijboer, T. C. W. (2011). The global effect: what determines where the eyes land? *Journal of eye movement research*, 4, 1-13.

Vigneau, F., Caissie, A.F. & Bors, D.A. (2006). Eye-movement analysis demonstrates strategic influences on intelligence. *Intelligence*, 34.261-272.

Wiedl, K. H. (1980). Kompensatorische Intervention in Rahmen intelligenzdiagnostischer Untersuchungen bei kognitiv impulsiven Kindern. [Compensatory intervention in intelligence assessment of cognitive impulsive children]. *Zeitschrift für Psychologie*, 9, 219-231.

Wishart, J. & Duffy, L. (1990). Instability of Performance on Cognitive Tests in Infants and Young Children with Down's Syndrome. *British Journal of Educational Psychology*, 60, 10-22.

- Yarbus, A. (1967). *Eye movement and vision*. New-York: Plenum Press. Jacobson, J.Z., & Dodwell, P.C. (1979). Saccadic eye movements during reading. *Brain and language*, 8, 303-314.
- Zambarbieri, D. (1992). Paradigmi di stimolazione e loro influenza sulle caratteristiche dei movimenti saccadici. In Traccis, S. et Zambarbieri, D. *I movimenti saccadici*. (pp. 61-81). Bologna: Patròn Editore.
- Zambarbieri, D. (2006). *Movimenti oculari*. Bologna: Patròn Editore.

V

ANNEXES

Annexe A : Protocole de passation du test Raven SPM (1998)

INSTRUCTIONS POUR L'ADMINISTRATION PAPIER-CRAYON DES SPM

Passation en auto-administration ou en collectif

Matériel

Il faut : des cahiers de test – ils peuvent être réutilisés à plusieurs reprises. On donne en outre à chaque sujet une feuille de réponse et un crayon. On peut se servir, pour donner les explications nécessaires, d'une reproduction (deux fois plus grande que l'originale) du problème A1 et de la feuille de réponse.

Les conditions d'examen

Le test peut être administré à un groupe de n'importe quelle taille, en fonction des conditions matérielles. Il faut compter environ une heure par groupe. Les sujets doivent être assis confortablement à des tables où ils auront la place de disposer leur cahier de test et leur feuille de réponses, et suffisamment espacées pour qu'ils ne puissent copier. Il faut que le superviseur ait la possibilité de passer aisément entre ces tables sans déranger les sujets. Ceux-ci doivent être tous assis confortablement à une table dans une pièce silencieuse, face à l'examinateur.

La procédure d'administration

On distribue les crayons et les feuilles de réponses. On prie les sujets d'inscrire sur la feuille de réponses leurs données biographiques (nom, prénom, date de naissance, etc.). Quand ils l'ont fait, on distribue les cahiers de tests, en demandant de ne pas les ouvrir avant que tout le monde soit prêt.

Dire : Ouvrez vos cahiers à la première page. Elle se présente ainsi.

Faire : Montrer au groupe un cahier ouvert, soit un agrandissement de démonstration.

Dire : En haut, vous voyez inscrit Série A, et ici, sur vos feuilles de réponses, vous avez une colonne pour la série A. Ceci est le problème A1. Vous voyez de quoi il s'agit ? La partie supérieure est un dessin dont il manque un morceau. Chacune des figures en dessous...

Faire : Les montrer l'une après l'autre.

Dire : ... a la forme qui convient pour combler l'espace, mais toutes ne complètent pas le dessin.

Sur le n°1...

Faire : Désigner la figure puis le grand dessin.

Dire : ...le dessin est complètement faux. Les n° 2 et 3 ne conviennent pas – ils comblent l'espace, mais ce n'est pas le bon dessin. Et le n°6 ? C'est le bon dessin...

Faire : Montrer que c'est le même que sur le pattern du haut.

Dire : ...mais il ne recouvre pas tout l'espace vide. Maintenant, posez votre doigt sur la figure qui convient tout à fait.

Faire : S'assurer que cela a été fait correctement. Au besoin, reprendre l'explication plus en détail puis

Dire : Oui, le n°4 est le bon. Donc, la réponse à A1 est 4 sur la feuille de réponses, écrivez 4 dans la case à côté de A1. Ne tournez pas encore la page.

Faire : Attendre que tout le monde ait fini et poursuivre :

Dire : Sur chaque page de votre cahier, il y a un dessin dont il manque un morceau. A chaque fois vous devez décider laquelle des figures en-dessous est celle qui convient pour compléter le grand dessin. Quand vous l'aurez trouvée, inscrivez son numéro sur votre feuille de réponses, à côté du numéro correspondant au grand dessin. N'écrivez rien dans vos cahiers car ils seront réutilisés avec d'autres élèves. Les problèmes sont simples au début, et deviennent plus difficiles au fur et à mesure que vous avancez. Il n'y a pas de piège. Si vous faites attention à la manière dont vous trouvez la solution des problèmes faciles, les suivants vous paraîtront moins difficiles. Essayez de les résoudre l'un après l'autre, en allant du début à la fin du cahier. Travaillez à votre rythme. Vous avez tout votre temps. Tournez la page et passez au problème suivant (A2). *Faites-le et ensuite on regarde si c'est bon pour tout le monde. Si c'est bon vous continuerez seuls.*

Faire : Quand il s'est écoulé suffisamment de temps pour que tout le monde ait inscrit sa réponse A2.

Dire : La bonne réponse est le n°5. Vérifiez que, sur votre feuille de réponses, vous avez bien écrit le chiffre 5 dans la case à côté de A2. Continuez comme ça, tout seuls, jusqu'à la fin du cahier. Je vais passer voir pour être sûre que vous avez compris les explications.

Supervision

Les sujets commettent souvent des erreurs en remplissant les feuilles de réponses. Les superviseurs doivent vérifier qu'ils y ont correctement noté leur réponse aux cinq premiers items. Un fois qu'un sujet a saisi la nature des premiers problèmes, le superviseur ne l'aide plus, en aucune façon, à saisir la méthode de raisonnement. Il se borne simplement à vérifier que les réponses sont reportées correctement. Il est fréquent que les sujets sautent un problème. Le superviseur doit s'assurer, quinze minutes après le début de l'épreuve, qu'ils inscrivent toujours la solution qu'ils proposent à un item en regard du numéro correspondant sur la feuille des réponses.

Au bout d'environ une demi-heure, on demande aux sujets d'indiquer quand ils ont fini. A ce moment-là, le superviseur vérifie que la feuille des réponses a été remplie correctement et que tous les problèmes ont été abordés. Puis on prie les sujets de rendre les cahiers et de sortir, ou de passer au test suivant s'il y a lieu.

Annexe B : Protocole de passation du test HART, version française

CONSIGNES POUR LA PASSATION DU TEST DE RAISONNEMENT ANALOGIQUE DE HESSELS (HART) – VERSION INFORMATIQUE

1. Présentation générale du test

Pour commencer, tu peux t'asseoir sur la chaise en face de l'ordinateur (en montrant la chaise en face de Tobii). Aujourd'hui, les exercices se passeront uniquement sur cet ordinateur (montrer Tobii). Tu vas faire des exercices pendant environ 10 minutes. Avant de commencer les exercices, nous allons faire quelques réglages. Une fois que nous aurons fait ces réglages, et pendant toute la durée des exercices, je te demande de ne pas tourner la tête et, si possible, de ne pas trop bouger.

Voilà comment les exercices vont se passer : pour chaque exercice, tu devras choisir ta réponse ; tu verras qu'il y a à chaque fois quatre possibilités, mais qu'il n'y a qu'une seule bonne réponse. A chaque fois, tu me diras le numéro de la réponse que tu as choisi et je noterai tes réponses sur une feuille. Entre chaque exercice, tu verras un rond gris apparaître au milieu de l'écran. Je te demande de bien regarder ce rond à chaque fois.

Je te demande de bien te concentrer et de bien réfléchir. Ça ne compte pas pour l'école, il n'y a pas de notes, ce n'est pas une évaluation mais essaye de faire le mieux possible. Au début c'est facile, mais après ça devient de plus en plus difficile. Tu ne trouveras pas toujours la bonne réponse; ce n'est pas grave. Ce qui m'intéresse, c'est de voir comment tu fais pour résoudre les exercices.

Grâce à « Show track status », vérifier la position de l'élève.

Débuter l'enregistrement et entrer les données personnelles de l'élève.

Items d'introduction

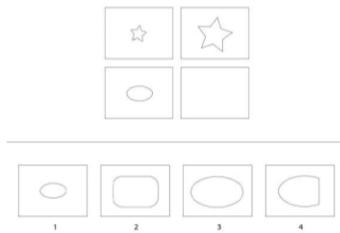
Nous allons faire ensemble un exemple pour que tu comprennes ce que je te demande.

Montrer l'exemple.

Tu vois un tableau avec deux colonnes et deux lignes (montrer les colonnes et les lignes). Toutes les cases sont remplies, sauf une. Là (montrer la case vide) il manque l'image et tu dois trouver celle qui va bien ici.

Alors dis-moi quelle image va bien dans la case vide (montrer la case vide) ?

Exemple Pré-test



Attendre la réponse de l'élève.

Tu as bien vu. C'est l'image 3. La 4, par exemple, est presque bonne, mais il manque un bout. Et l'image 2 est totalement fausse.

Maintenant on peut continuer. N'oublie pas que, pendant toute la durée des exercices, il ne faut pas tourner la tête et, si possible, de ne pas trop bouger.

2. Entraînement

Nous allons faire quelques exercices ensemble pour que tu comprennes comment bien résoudre les exercices que tu as faits.

Mettre la figure 1 à l'écran.

Figure 1



Tu as une première figure ici (montrer le premier rectangle), j'aimerais que tu la compares avec les autres figures (montrer les 5 autres figures). Tu peux me décrire chaque figure en me disant ce qui change et ce qui reste pareil?

Attendre la réponse de l'élève.

Oui c'est très bien. A gauche il y a un rectangle blanc placé verticalement (montrez le rectangle), à droite il y a aussi des rectangles mais parfois c'est l'orientation qui change (montrez le 1er rectangle), parfois la couleur (montrez le 2ème et 4ème) et même une fois l'épaisseur du trait qui change (montrez la 5ème). Quand il y a plusieurs images, il faut toujours bien regarder ce qui change ou ce qui reste identique en comparant les figures entre elles.

Mettre la figure 2 à l'écran.

Figure 2



Tu peux me décrire chaque figure en me disant ce qui change et ce qui reste pareil?

Attendre la réponse de l'élève.

Oui c'est très bien. A gauche il y a un rectangle gris placé verticalement (montrez le rectangle), à droite il y a aussi des rectangles, ils sont tous identiques à celui de gauche, sauf le 2ème qui n'a pas de bordure noire (montrez le 2ème). Il faut bien faire attention car des fois il y a de toutes petites choses qui changent.

Procéder de la même manière pour les deux figures suivantes et dire à la fin de l'exercice :

Dans cet exercice, on a vu qu'il est très important de comparer les figures entre elles pour voir ce qui change ou ce qui reste identique. Il faut toujours bien faire attention aux petits détails, car lorsque deux figures se ressemblent beaucoup il y a souvent un petit détail qui change

Figure 3

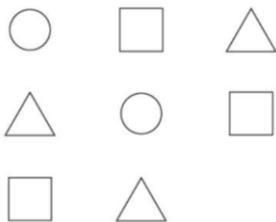


Figure 4



Mettre la figure 5 à l'écran.

Figure 5



Peux-tu me décrire les changements que tu vois dans cette figure? Attendre la réponse de l'élève.

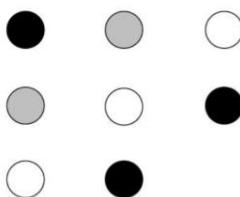
Oui c'est très bien. Sur la première ligne, il y a un rond, un carré puis un triangle et sur la deuxième ligne il y a d'abord un triangle puis un rond puis un carré. Il y a donc trois formes différentes sur chaque ligne mais elles sont toutes de la même couleur.

Peux-tu me décrire la figure que l'on doit mettre ici (Montrer la case vide) et m'expliquer pourquoi ?

Oui c'est juste. Il faut un rond car sur chaque ligne il y a toujours les trois mêmes figures (un rond, un carré et un triangle) mais elles changent de place. Elles bougent en allant vers la droite, chaque figure est une fois à gauche, une fois au milieu et une fois à droite (montrer le rond qui se déplace).

Mettre la figure 6 à l'écran.

Figure 6



Peux-tu me décrire les changements que tu vois dans cette figure? Attendre la réponse de l'élève.

Oui c'est très bien. Sur la première ligne il y a un rond noir puis un rond gris puis un rond blanc. Sur la deuxième ligne, il y a d'abord un rond gris, puis un rond blanc puis un noir. Sur chaque ligne, les trois figures ont la même forme mais la couleur change.

Peux-tu me décrire la figure que l'on doit mettre ici (Montrer la case vide) et m'expliquer pourquoi ?

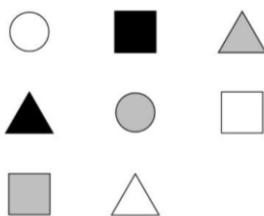
Oui c'est très bien. Il faut un rond gris car sur chaque ligne il y a chaque fois trois ronds avec les trois mêmes couleurs mais leur place change. Les couleurs bougent en allant vers la gauche. Chaque couleur est une fois à droite, une fois au milieu et une fois à gauche (montrer le rond blanc qui se déplace vers la gauche).

Mettre la figure 7 à l'écran.

Peux-tu me décrire les changements que tu vois dans cette figure ? Attendre la réponse de l'élève.

Oui c'est très bien. Sur la première ligne il y a un rond blanc puis un carré noir puis un triangle gris. Sur la deuxième ligne, il y a d'abord un triangle noir puis un rond gris puis un carré blanc. Sur chaque ligne, il y a trois figures différentes et trois couleurs différentes.

Figure 7

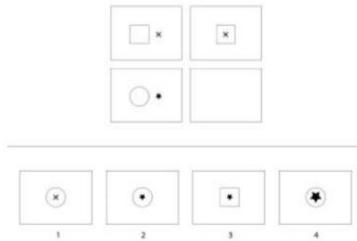


Peux-tu me décrire la figure que l'on doit mettre ici (Montrer la case vide) et m'expliquer pourquoi ?

Oui c'est juste. Il faut un rond noir car il y a deux choses qui changent: la couleur et la forme des figures. Les couleurs se déplacent vers la gauche et les formes vers la droite. Pour vérifier on peut voir que chaque forme et chaque couleur se trouve une fois à droite, une fois au milieu et une fois à gauche (montrer le rond qui se déplace vers la droite puis la couleur blanche qui se déplace vers la gauche.).

On va maintenant faire ensemble six exercices en utilisant ce que l'on vient d'apprendre.

Exercice 1



Mettre l'exercice 1 à l'écran et ne montrer que la première ligne de la matrice. Cacher le reste.

Peux-tu décrire ces deux images en me disant ce qui change ?

Oui, tu as bien décrit. A gauche, il y a un carré blanc avec une croix à côté du carré. A droite, il y a aussi un carré blanc, mais cette fois la croix se trouve à l'intérieur du carré.

Montrer la deuxième ligne en cachant les alternatives.

Peux-tu décrire cette image (montrer M3) et me dire ce qui change par rapport à celle du dessus (montrer M1) ?

Attendre la réponse de l'élève.

Sur la deuxième ligne, on a à gauche un rond blanc avec une étoile à côté. Cette fois, les formes sont différentes. On a un rond et une étoile. Donc de haut en bas, c'est la forme qui change (Accompagner l'explication en montrant les formes de la matrice).

Pour trouver l'image qui va bien en bas à droite il faut faire le même changement que dans la première ligne.

Imagine-toi dans ta tête l'image qui va bien dans la case vide (montrer la case).

Peux-tu me décrire l'image qui va bien ici ? Attendre la réponse de l'élève.

Montrer les 4 choix de réponse.

Maintenant, tu peux comparer l'image que tu as dans ta tête avec chacune des quatre alternatives et choisir la bonne réponse. Fais attention aux petits détails, il n'y a toujours qu'une seule bonne réponse.

Est-ce que c'est la 1, pourquoi ? Attendre la réponse de l'élève.

Non, parce qu'il s'agit bien d'un rond mais c'est une étoile qui va avec et pas une croix. Est-ce

que c'est la 2, pourquoi ?

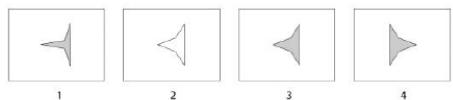
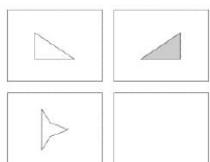
Attendre la réponse de l'élève.

Oui c'est bien la réponse 2. De gauche à droite (montrer M1 puis M2) on voit que la petite forme est d'abord à côté de la grande, puis à l'intérieur de la grande forme. De haut en bas (montrer M1 puis M3) c'est la forme qui change, on a d'abord un carré et une croix, puis un rond et une étoile. Donc on fait la même chose ici (montrer M2), on cherche un rond avec une étoile à l'intérieur.

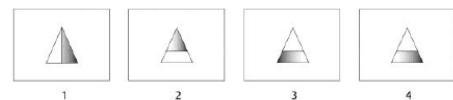
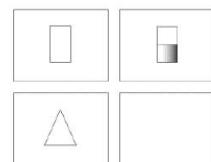
Exercice 2 – Exercice 6

Procéder de la même manière que lors du premier exercice.

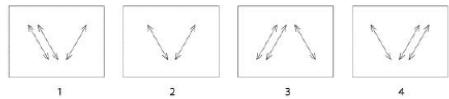
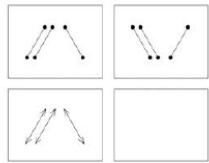
Exercice 2



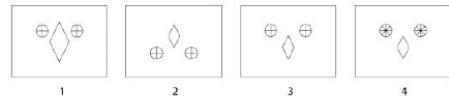
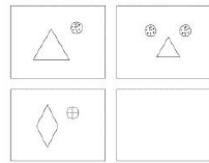
Exercice 3



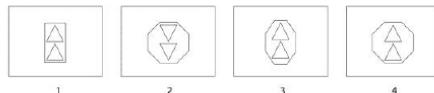
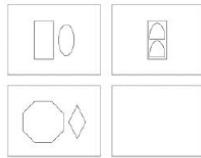
Exercice 4



Exercice 5



Exercice 6



3. Post-test

Pour celles et ceux qui ont fait l'entraînement :

Maintenant que tu as fait quelques exercices pour t'entraîner, tu vas faire une autre série d'exercices.

Pour celles et ceux qui n'ont pas fait l'entraînement :

Aujourd'hui, nous allons faire une autre série d'exercices.

Pour tout le monde :

Comme la dernière fois, on commence par faire quelques réglages. Une fois que nous aurons fait ces réglages, et pendant toute la durée des exercices, je te demande de ne pas tourner la tête et, si possible, de ne pas trop bouger.

Comme lors de la première série d'exercices, il y a aura à nouveau trois images, et tu devras trouver la quatrième image qui complète bien la série. Tu auras quatre propositions, mais une seule réponse est correcte.

Comme la dernière fois, pour chaque exercice, tu me diras le numéro de la réponse que tu auras choisi et je noterai tes réponses sur une feuille.

Grâce à « Show track status », vérifier la position de l'élève.

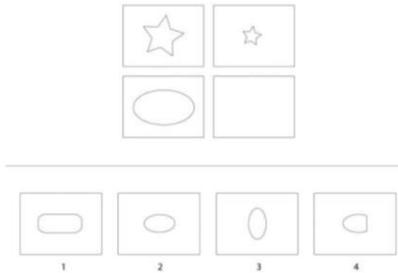
Débuter l'enregistrement et entrer les données personnelles de l'élève.

Items d'introduction

Nous allons faire ensemble un exemple avant que tu commences.

Montrer l'exemple.

Exemple Post-test



C'est comme la dernière fois. Tu vois un tableau avec deux colonnes et deux lignes (montrer les colonnes et les lignes). Toutes les cases sont remplies, sauf une. Là (montrer la case vide) il manque l'image et tu dois trouver celle qui va bien ici.

Alors dis-moi quelle image va bien dans la case vide (montrer la case vide) ?

Attendre la réponse de l'élève.

Tu as bien vu. C'est l'image 2. La 4, par exemple, est presque bonne, mais il manque un bout. Et l'image 1 est totalement fausse.

Maintenant on peut continuer. N'oublie pas que, pendant toute la durée des exercices, il ne faut pas tourner la tête et, si possible, de ne pas trop bouger.

Annexe C : Protocole de passation du test HART, version italienne

1. Presentazione del test

Per cominciare puoi sederti su questa sedia davanti al computer (*mostrare la sedia di fronte a Tobii*). Gli esercizi che faremo oggi si svolgeranno su questo computer (*mostrare Tobii*). Dovrai fare degli esercizi per circa una decina di minuti. Prima di cominciare devo assicurarmi che tutto funzioni bene. Per piacere, cerca di non girare la testa e di non muoverti troppo.

Ecco come si svolgeranno gli esercizi: per ogni esercizio dovrai scegliere la tua risposta; come vedrai in seguito, ognuno di essi propone quattro risposte: solamente una è quella corretta. Ogni volta dovrai dirmi il numero della risposta che hai scelto: io le scriverò su un foglio. Tra un esercizio e l'altro vedrai apparire un cerchio grigio al centro dello schermo. È molto importante che tu lo guardi ogni volta.

Ti chiedo per piacere di cercare di concentrarti e di riflettere bene. Ricordati che gli esercizi di oggi non sono legati alla scuola: non ci sono voti e non si tratta di un test. Ma tu cerca di fare il meglio possibile. All'inizio gli esercizi saranno facili, ma in seguito il livello di difficoltà andrà crescendo. Non troverai sempre la buona soluzione, non è grave. Quello che mi interessa è vedere quale metodo adotterai per risolvere gli esercizi.

Con l'aiuto di "Show track status" verificare la posizione dell'allievo. Iniziare la registrazione ed entrare nei dati personali dell'allievo.

2. Items di introduzione

Adesso faremo insieme qualche esempio, in modo che tu possa capire quello che ti domando di fare negli esercizi.

Mostrare l'esempio.

In questa tabella ci sono due colonne e due linee (*mostrare le colonne e le linee*). Tutte le caselle sono piene, salvo una. In questa (*mostrare la casella vuota*) manca l'immagine, sarai tu a trovare quella giusta.

Allora, prova a dirmi quale immagine potrebbe andare bene in questa casella vuota (*mostrare la casella vuota*).

Molto bene, è l'immagine 3. L'immagine 4 è quasi giusta ma manca un pezzetto. L'immagine 2 invece è completamente sbagliata.

Adesso possiamo cominciare. Per piacere, ricordati di non girare la testa e di restare il più fermo possibile.

3. Allenamento

Ora faremo insieme qualche esercizio affinché tu possa capire come bisogna risolvere gli esercizi.

Mettere la figura 1 sullo schermo.

Hai una prima figura qui (*mostrare il primo rettangolo*), mi piacerebbe che tu la comparassi con le altre figure (*mostrare le 5 altre figure*). Potresti descrivermi ognuna di esse dicendomi che cosa cambia e che cosa resta uguale?

Aspettare la risposta dell'allievo.

Sì, molto bene. A sinistra c'è un rettangolo bianco in posizione verticale (*mostrare il rettangolo*), anche alla sua destra ci sono dei rettangoli, ma questi sono diversi talvolta per l'orientamento (*mostrare il primo rettangolo*), talvolta per il colore (*mostrare il secondo e il quarto*) e in questo caso (*mostrare il quinto*) per lo spessore, che è differente.

Mettete la figura 2 allo schermo.

Potresti descrivermi questa figura e dirmi che cosa cambia e che cosa invece resta uguale?

Aspettare la risposta dell'allievo.

Sì, molto bene. A sinistra c'è un rettangolo grigio posizionato in maniera verticale (*mostrare il rettangolo*), alla sua destra ci sono degli altri rettangoli che sono tutti identici al primo.

Solamente il secondo non ha lo stesso bordo nero (*mostrare il secondo*). Bisogna fare molta attenzione, siccome in alcuni casi ci sono dei piccoli dettagli che cambiano.

Procedere nella stessa maniera per le due figure seguenti e dire alla fine dell'esercizio:

In questo esercizio abbiamo visto che è molto importante confrontare le figure tra di loro per capire che cosa cambia o che cosa resta identico. Bisogna sempre fare bene attenzione ai piccoli

dettagli perché spesso, quando due figure si assomigliano molto, c'è solo un minimo particolare che fa la differenza.

Mettere la figura 5 sullo schermo.

Potresti descrivermi le differenze che vedi in questa immagine?

Aspettare la risposta dell'allievo.

Bene, bravo/a. Sulla prima linea ci sono un cerchio, un quadrato e un triangolo. Sulla seconda linea invece c'è prima un triangolo, poi un cerchio e per finire un quadrato. Ci sono quindi tre forme differenti su ogni linea, ma sono tutte dello stesso colore.

Potresti descrivermi la figura che bisognerebbe mettere qui (*mostrare lo spazio vuoto*) e spiegarmi come mai?

Sì, è giusto. La figura che ci serve è un cerchio siccome su ogni linea ci sono sempre le stesse tre figure (un cerchio, un quadrato e un triangolo) che cambiano posizione su ogni linea. Le figure si muovono andando verso destra, ogni figura è una volta a destra, una volta al centro e una volta a sinistra (*mostrare il rotondo che si sposta*).

Mettere la figura 6 allo schermo.

Potresti descrivermi le differenze che ci sono in questa figura?

Aspettare la risposta dall'allievo.

Bene, bravo/a. Sulla prima linea ci sono un cerchio nero, poi uno grigio e per finire uno bianco. Sulla seconda linea, troviamo in prima posizione un cerchio grigio, poi uno bianco e infine uno nero. Su ogni linea troviamo tre figure identiche, ma di colori diversi.

Potresti descrivermi la figura che bisognerebbe mettere qui (*mostrare lo spazio vuoto*) e spiegarmi come mai?

Sì, molto bene. La figura che ci serve è un cerchio grigio siccome su ogni linea ci sono ogni volta tre cerchi che hanno gli stessi colori. Quello che cambia è la loro posizione. I colori si spostano andando verso sinistra. Ogni colore è una volta a destra, una volta al centro e una volta a sinistra (*mostrare il rotondo bianco che si sposta verso sinistra*).

Mettete la figura 7 sullo schermo.

Potresti descrivermi i cambiamenti che vedi in questa immagine?

Aspettare la risposta dell'allievo.

Bravo/a, molto bene. Sulla prima linea ci sono un cerchio bianco, un quadrato nero e un triangolo grigio. Nella seconda linea troviamo prima un triangolo nero, poi un cerchio grigio e infine un quadrato bianco. Su ogni linea ci sono tre figure diverse e tre colori diversi.

Potresti descrivermi la figura che dobbiamo mettere qui (*mostrare lo spazio vuoto*) e spiegarmi come mai?

Sì, è giusto. La figura che ci serve è un cerchio nero siccome sono presenti due caratteristiche che cambiano: il colore e la forma delle figure. I colori si spostano verso sinistra e le forme verso destra. Per verificarlo possiamo vedere che ogni forma e ogni colore si trovano una volta a destra, una volta al centro e una volta a sinistra (mostrare il cerchio che si sposta verso destra, poi mostrare il colore bianco che si sposta verso sinistra).

Molto bene, ora faremo insieme sei esercizi utilizzando quello che abbiamo appena imparato.

Mettere l'esercizio 1 sullo schermo e mostrare solamente la prima linea della matrice.

Nascondere il resto.

Potresti descrivermi queste due immagini dicendomi che cosa cambia?

Bene, buona descrizione. A sinistra c'è un quadrato bianco con accanto una croce. Anche a destra c'è un quadrato bianco, ma questa volta la croce si trova al suo interno.

Mostrare la seconda linea e nascondere le alternative.

Potresti descrivermi questa immagine (*mostrare M3*) e dirmi che cosa cambia rispetto a quella vista sopra (*mostrare M1*)?

Aspettare la risposta dell'allievo.

Sulla seconda linea, a sinistra, c'è un cerchio bianco con accanto una stella. Questa volta le forme sono diverse. Ci sono un cerchio ed una stella. Quindi, dall'alto al basso, è la forma che cambia (*accompagnare la spiegazione mostrando le forme della matrice*).

Per trovare la figura corretta da inserire in fondo a destra bisogna fare lo stesso cambiamento che abbiamo osservato nella prima linea.

Immagina nella tua testa la figura che potrebbe andare bene nella casella vuota (*mostrare la casella*).

Potresti descrivermi la figura che va bene qui?

Attendere la risposta dell'allievo.

Mostrare le 4 alternative di risposta.

Ora puoi paragonare la figura che ti sei immaginato con ognuna di queste quattro alternative e scegliere la risposta corretta. Fai attenzione ai piccoli dettagli. C'è sempre una sola risposta corretta.

Hai scelto la prima? Perché?

Aspettare la risposta dell'allievo.

No, hai ragione: si tratta di un cerchio, ma dovrebbe esserci una stella e non una croce.

Potrebbe essere la seconda? Perché?

Aspettare la risposta dell'allievo.

Bene, è proprio la seconda. Da sinistra a destra (*mostrare M1, poi M2*) si vede bene che la piccola forma si trova prima accanto alla grande e poi all'interno di essa. Dall'alto al basso (*mostrare M1, poi M3*) è la forma che cambia. Abbiamo prima un quadrato con una croce, poi un cerchio con una stella. Quindi, per trovare la buona soluzione, dobbiamo procedere nello stesso modo anche qui (*mostrare M2*): cerchiamo un cerchio con una stella al suo interno.

Procedere nella medesima maniera utilizzata per il primo esercizio.

4. Post-test

Per gli allievi che hanno fatto l'allenamento:

Ora che hai fatto qualche esercizio per allenarti, dovrà fare un'altra serie di esercizi.

Per gli allievi che non hanno fatto l'allenamento:

Oggi faremo un'altra serie di esercizi.

Per tutti gli allievi:

Come l'ultima volta, devo assicurarmi che tutto funzioni bene. Per piacere, ricordati di non girare la testa e cerca di non muoverti troppo.

Come è già capitato durante la prima serie di esercizi, ci saranno di nuovo tre figure e dovrà trovare la quarta che completa correttamente la serie. Avrai a disposizione quattro proposte, ma solamente una è quella corretta.

Come l'ultima volta, alla fine di ogni esercizio, dovrà dirmi il numero della risposta che hai scelto e io le scriverò su un foglio.

Con l'aiuto di "Show track status" verificare la posizione dell'allievo.

Cominciare la registrazione ed entrare nei dati personali dell'allievo.

5. Items di introduzione

Allora, prima di cominciare faremo insieme un esempio.

Mostrare l'esempio.

È come l'altra volta. C'è una tabella con due colonne e due linee (*mostrare le colonne e le linee*).

Tutte le caselle sono piene, salvo una. In questa (*mostrare la casella vuota*) manca una figura.

Sei tu che dovrà trovare quella giusta da inserire.

Allora, secondo te, quale immagine bisognerebbe inserire nella casella vuota (*mostrare la casella vuota*)?

Aspettare la risposta dell'allievo.

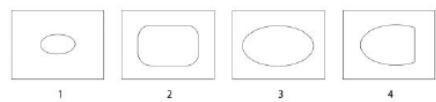
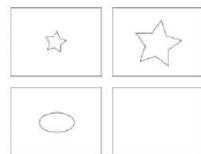
Molto bene, è l'immagine 2. L'immagine 4 è quasi giusta, ma come vedi manca un pezzetto.

L'immagine 1 invece è completamente scorretta.

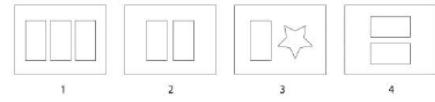
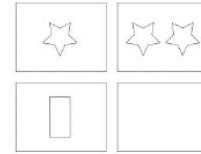
Adesso possiamo continuare. Per piacere, ricordati di non girare la testa e cerca di muoverti il meno possibile.

Annexe D : Items utilisés pour le pré-test du HART

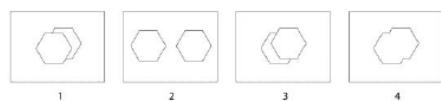
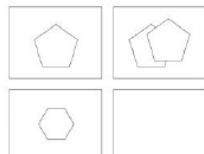
Exemple



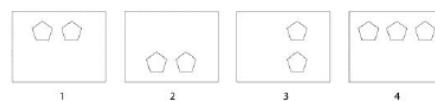
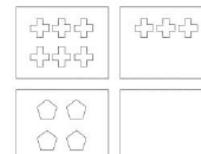
Item 1



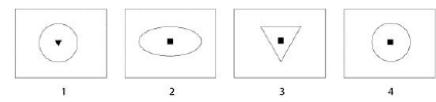
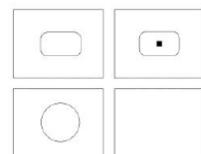
Item 2



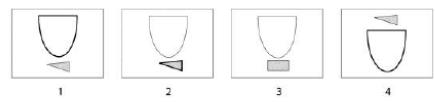
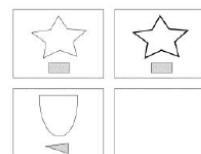
Item 3



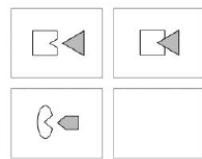
Item 4



Item 5

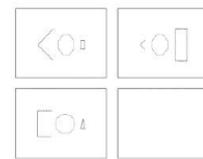


Item 6



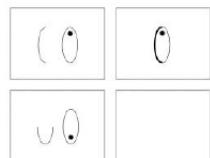
-
- 1 2 3 4

Item 7



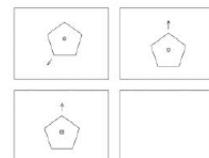
-
- 1 2 3 4

Item 8



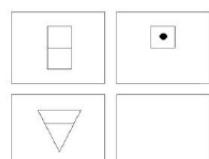
-
- 1 2 3 4

Item 9



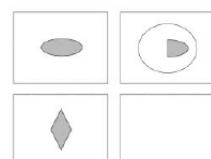
-
- 1 2 3 4

Item 10



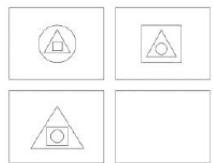
-
- 1 2 3 4

Item 11



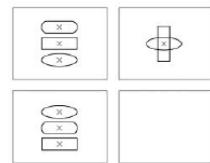
-
- 1 2 3 4

Item 12



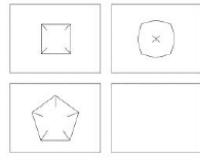
-
- | | | | |
|--|--|--|--|
| | | | |
|--|--|--|--|

Item 13



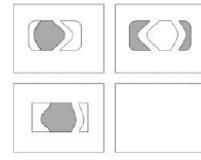
-
- | | | | |
|--|--|--|--|
| | | | |
|--|--|--|--|

Item 14



-
- | | | | |
|--|--|--|--|
| | | | |
|--|--|--|--|

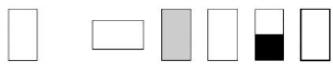
Item 15



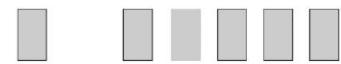
-
- | | | | |
|--|--|--|--|
| | | | |
|--|--|--|--|

Annexe E : Exercices et items utilisés pour l'entraînement

Exercice comparaison 1



Exercice comparaison 2



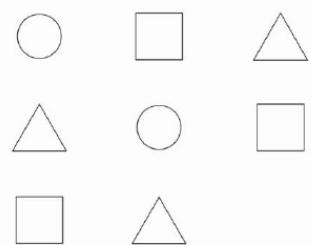
Exercice comparaison 3



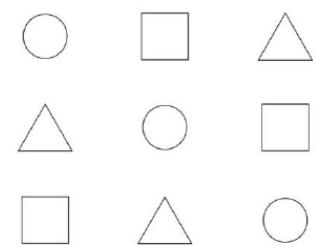
Exercice comparaison 4



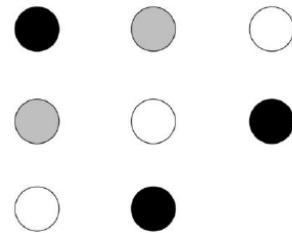
Exercice inférence 5



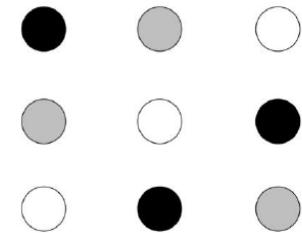
Réponse inférence 5



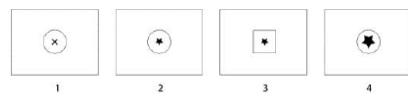
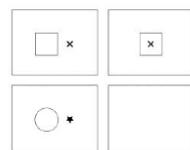
Exercice inférence 6



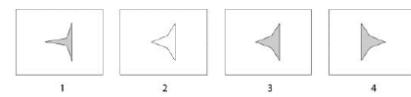
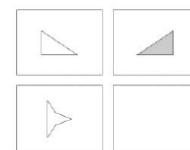
Réponse inférence 6



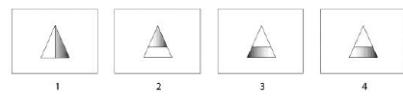
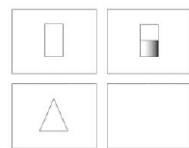
Item 1



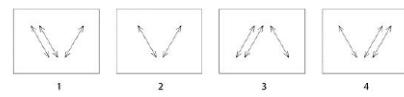
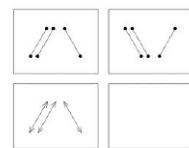
Item 2



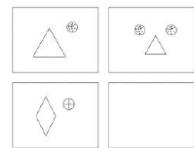
Item 3



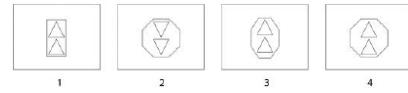
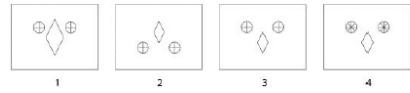
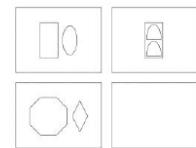
Item 4



Item 5

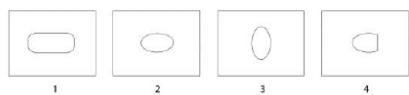
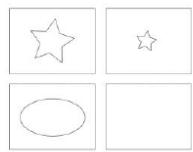


Item 6

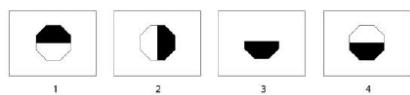
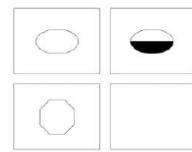


Annexe F : Items utilisés pour le post-test

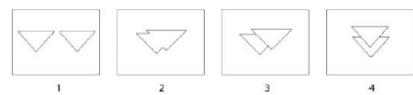
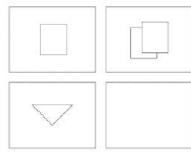
Exemple



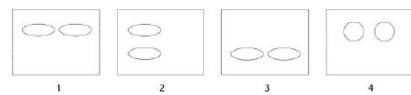
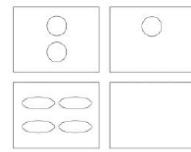
Item 1



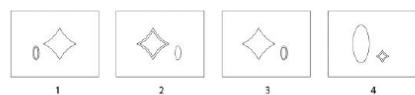
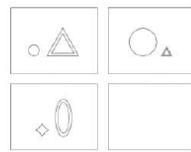
Item 2



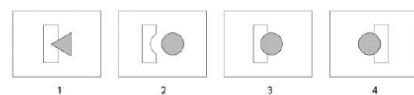
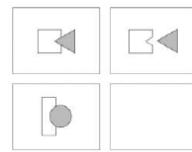
Item 3



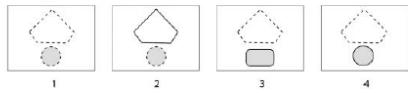
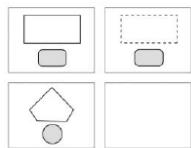
Item 4



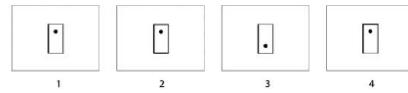
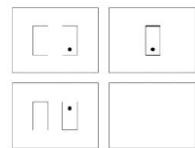
Item 5



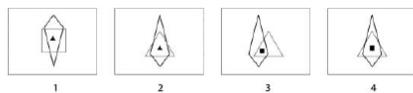
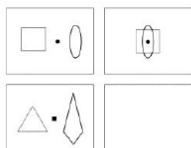
Item 6



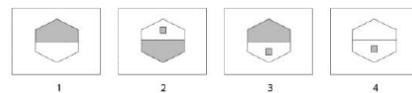
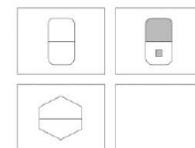
Item 7



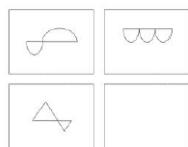
Item 8



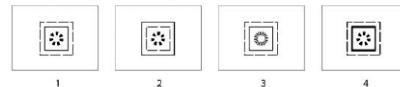
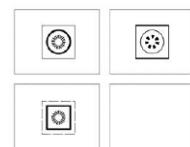
Item 9



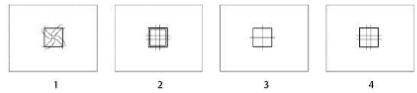
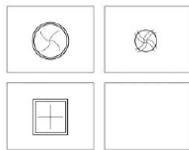
Item 10



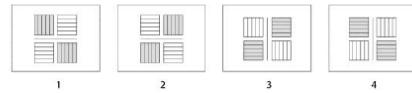
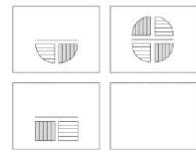
Item 11



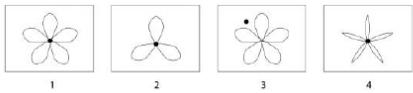
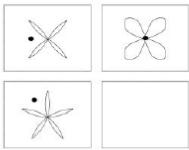
Item 12



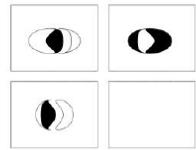
Item 13



Item 14

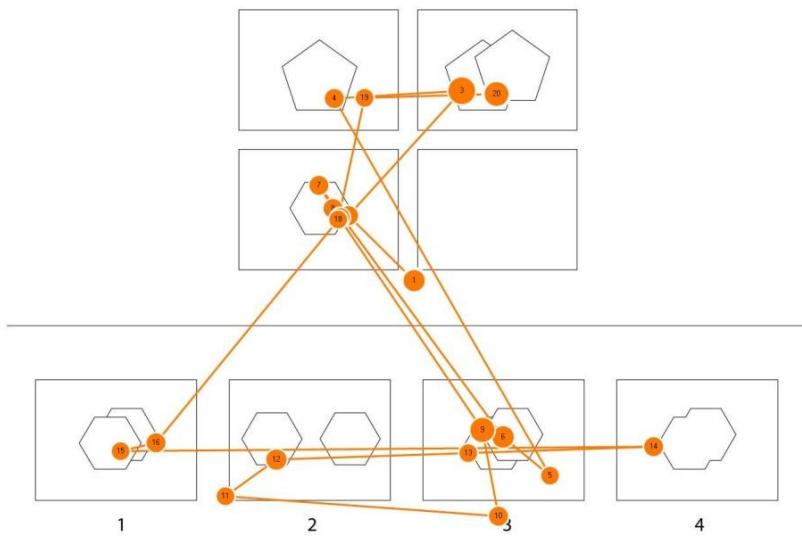


Item 15

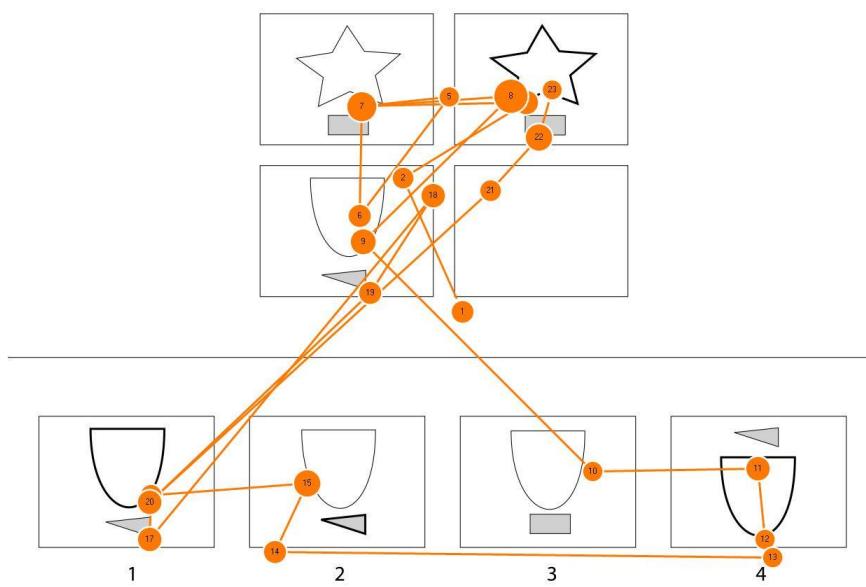


Annexe G : Exemples d'images découlant de l'outil Gaze Plot

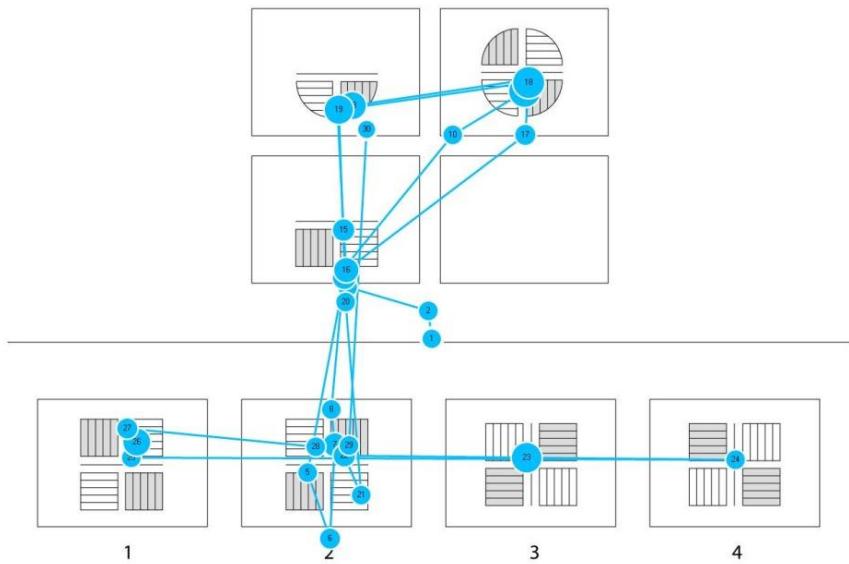
Media: 2 N-EYE-45.jpg
Time: 00:00:00.000 - 00:00:30.178
Participant filter: All



Media: 5 N-EYE-52.jpg
Time: 00:00:00.000 - 00:00:26.698
Participant filter: All



Media: 13.N-EYEP-G1.jpg
Time: 00:00:00.000 - 00:00:33.023
Participant filter: All



Media: 3.N-EYEP-A10.jpg
Time: 00:00:00.000 - 00:00:59.858
Participant filter: All

