

# PERCEVOIR LA STRUCTURE D'UNE SITUATION: UNE APPROCHE DIFFÉRENTIELLE

---

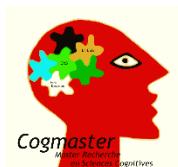
PRÉSENTÉE PAR LAURE ZEBOULON

SOUS LA CO-DIRECTION D' E. PASQUINELLI & E. SANDER

---

MÉMOIRE DE RECHERCHE EN SCIENCES COGNITIVES  
MASTER 2 - COGMASTER

ANNEE UNIVERSITAIRE 2017-2018



## Nombre de mots

---

Nombre de mots du présent mémoire (hors résumé, table des matières, page de garde, titres, déclarations, remerciements, bibliographie et annexes)

= 14 346 mots

## Remerciements

---

Je tiens à exprimer mes plus sincères remerciements à Elena Pasquinelli et Emmanuel Sander non seulement pour avoir assuré la direction de ce mémoire et m'avoir donné l'opportunité de travailler sur ce sujet passionnant, mais aussi et surtout pour leur extraordinaire gentillesse et soutien tout au long de ces deux années. L'intelligence de leurs réflexions, la pertinence de leurs conseils, leurs passions n'ont cessé d'être des permanentes sources d'inspiration.

J'exprime ma profonde gratitude aux équipes administratives et pédagogiques du Cogmaster pour m'avoir donné la chance de suivre ce master dans les meilleures conditions possibles et pour la qualité de l'enseignement dispensé. J'adresse des remerciements particuliers aux professeurs et tuteurs Jean-Pierre Nadal, Roberto Cassati pour leur bienveillance.

Je remercie l'ensemble des collaborateurs de la Map pour leur accueil et tout particulièrement Muriel Treill, Fatima Rahmoun et Katia Allegraud pour leur aide dans l'organisation des passations dans les collèges.

Un grand merci aux professeurs, aux chefs d'établissement ainsi qu'aux 406 élèves de 6eme, qui ont participé à cette étude. Ces moments passés dans les classes ont été une expérience très enrichissante.

Je remercie chaleureusement Katarina Gvozdic et Hippolyte Gros pour leurs précieux conseils.

Last but not least, je tiens à exprimer toute ma gratitude et mon amour à ma famille pour leur soutien ces deux dernières années.

## Déclaration d'originalité

---

Le présent mémoire porte sur l'évaluation chez des élèves de 6ème de leur capacité à percevoir des structures communes à des situations au-delà de l'habillage de surface. S'inscrivant dans la lignée des travaux s'intéressant à l'influence des différents types de similarité dans le transfert (Gick et Holyoak, 1983; Gentner et al , 2003; Raynal, Clément et Sander ; 2017) et de ceux étudiant l'organisation conceptuelle des connaissances des novices et des experts dans certaines matières scientifiques (Chi, Feltovich et Glaser, 1981 ; Smith et al, 2013 ; Krieter, Julius, Tanner, Bush et Scott, 2016), cette étude s'en distingue par le choix de son paradigme expérimental et des stimuli utilisés.

Notre protocole expérimental s'articule autour de tâches de catégorisation et de sélection d'analogie basées sur des récits de la vie quotidienne, ne faisant appel à aucune connaissance scolaire spécifique et dont les concepts les plus abstraits sont connus de tous. Si des tâches de rappel et de production d'analogie se sont basés sur ces petits récits de la vie quotidienne, des tâches de catégorisation n'ont, à notre connaissance, jamais été réalisées sur ce type d'énoncés et se sont focalisées jusqu'alors sur des problèmes scientifiques (tel ceux de physique, biologie ou chimie). L'ensemble des textes ainsi que le design expérimental (structure des différentes tâches, matériel utilisé) ont été spécialement élaborés pour cette étude.

## Déclaration de contribution

---

Les principaux contributeurs à ce présent mémoire ont été Elena Pasquinelli (EP) et Emmanuel Sander (ES) qui l'ont codirigé, Laure Zeboulon (LZ), les collaborateurs de la Fondation la Map : Muriel Treill (MT), Fatima Rahmoun (FR), Katia Allegraud (KA) ainsi que les professeurs : Aube Mangin (AM), Amandine Long (AL), Manuelle Zeitoun (MZ) et les directions des différents établissements qui ont permis la passation des expériences.

Plus précisément les contributions ont été les suivantes :

- Idée initiale de l'utilisation de petits récits de la vie quotidienne comme stimuli pour les tâches d'analogie : ES.
- Idée initiale des tâches de catégorisation sous forme de jeu de cartes à trier : EP.
- Définition de l'approche générale et des hypothèses : EP, ES, LZ.
- Revue littéraire : LZ.
- Approche méthodologique : EP, ES, LZ
- Élaboration du protocole expérimental : ES, LZ.
- Introduction auprès des collègues : EP, MT, FR, KA.
- Sélection des classes : LZ.
- Création des énoncés : ES, LZ.
- Création du matériel expérimental : LZ.
- Passation des expériences : LZ (en présence des professeurs).
- Codage des résultats : LZ.
- Analyse des données : LZ.
- Interprétation des résultats et formulation des conclusions : ES., LZ.
- Rédaction du mémoire : LZ.
- Relecture, corrections et conseils : EP, ES.

## **TABLE DES MATIERES**

<b>I.</b>	<b>Introduction .....</b>	<b>9</b>
<b>II.</b>	<b>Cadre théorique .....</b>	<b>11</b>
A.	Comment percevons-nous le monde ? .....	11
1.	Théorie classique des concepts .....	11
2.	Théorie des prototypes & Théorie des exemplaires .....	11
3.	Théorie des connaissances .....	14
B.	Comment catégorisons-nous les situations ? .....	15
1.	Modèles de compréhension de texte.....	15
a)	Modèles des schémas .....	16
b)	Modèles de situation .....	16
2.	Analogie .....	17
C.	Comment développer notre esprit critique ? .....	20
1.	Expertise .....	20
2.	Esprit Critique .....	21
<b>III.</b>	<b>Problématique et hypothèses .....</b>	<b>23</b>
<b>IV.</b>	<b>Méthodologie .....</b>	<b>26</b>
A.	Participants .....	26
B.	Matériel .....	28
C.	Procédure .....	36
D.	Analyse quantitative .....	35
<b>V.</b>	<b>Résultats .....</b>	<b>41</b>
A.	Tâche 1 de catégorisation .....	42
1.	Analyses préliminaires .....	41
2.	Analyses principales .....	45
B.	Tâche 2 de détection d'analogie.....	51

1. Analyses préliminaires .....	52
2. Analyses principales .....	53
<b>VI. Discussion .....</b>	<b>57</b>
<b>VII. Conclusion .....</b>	<b>61</b>
<b>VIII. Bibliographie .....</b>	<b>64</b>
<b>IX. Annexes .....</b>	<b>69</b>

## I. Introduction

---

L’analogie est au cœur de la cognition humaine (Hofstadter, 2011) et des théories générales sur l’intelligence (Spearman, 1923 ; Sternberg, 1977). L’analogie permet d’aborder l’inconnu à partir de ce que l’on connaît (Sander, 2000). C’est un outil puissant de compréhension, de raisonnement et d’apprentissage parce qu’elle permet de faire des inférences à partir de connaissances acquises ou d’expériences vécues.

Raisonner par analogie, faire preuve de flexibilité cognitive nécessitent de percevoir les différentes structures sous-jacentes afin d’en saisir les relations de similitude au-delà des éléments superficiels. La perception d’une structure profonde pertinente à une situation marque une compréhension fine de celle-ci et peut être considérée comme une « intelligence » des choses qui s’inscrit dans le cadre de l’appréciation critique du monde qui nous entoure. En effet, toucher l’essence même d’un problème, analyser une situation sous différents angles, ne pas céder aux catégorisations hâtives représentent des enjeux éducatifs majeurs qui constituent des bases du développement de l’esprit critique.

De nombreuses recherches expérimentales portent sur le transfert analogique et ses conditions de réalisation, ainsi que sur le rôle des différents types de similarité dans les mécanismes élémentaires d’évocation ou d’appariement qui contribuent au transfert. Un nombre plus restreint d’études s’est concentré sur la capacité à percevoir des similitudes à différents niveaux d’abstraction et sur l’organisation conceptuelle des connaissances qu’elle révèle. Parmi eux, les travaux de Chi, Feltovitch et Glaser (1981) ont apporté une contribution majeure. En effet, en s’appuyant sur des tâches de catégorisation d’exercices de physique ils ont montré que les experts diffèrent des novices par leur tendance à utiliser des critères de classification basés sur les éléments structurels plutôt que superficiels.

Sur la base de ces travaux, plusieurs études portant sur des tâches de catégorisation d’énoncés de problèmes scientifiques ont été menées ces

dernières années dans la perspective de développer des outils pédagogiques permettant de mettre en évidence non plus l'acquisition de certaines connaissances spécifiques mais l'organisation conceptuelle de ces connaissances en la comparant avec celle d'experts dans le domaine.

Afin de répondre à nos questions de recherche, nous avons souhaité appliquer cette approche non pas à des problèmes scientifiques, mais à des petits récits de la vie quotidienne qui ne font appel à aucune connaissance dans un domaine particulier. L'objectif était de mettre en évidence les différences de représentations mentales et de degrés d'abstraction des structures perçues par les élèves dans des situations où les concepts abstraits sont connus de tous.

En nous intéressant à différentes populations d'élèves de 6ème, nous avons souhaité répondre aux questions scientifiques suivantes : comment la capacité de percevoir les structures sous-jacentes à une situation varie-t-elle d'un élève à un autre ? Est-elle discriminante entre différentes populations d'enfants ? Ces tâches de catégorisation de textes peuvent-ils constituer des outils pédagogiques permettant de distinguer les différentes perçues par les élèves en fonction de niveau d'abstraction ?

## II. Cadre Théorique

---

### A. COMMENT PERCEVONS-NOUS LE MONDE ?

Chaque chose, chaque situation rencontrée par un individu est appréhendée par le biais de concepts. Les concepts incarnent notre connaissance des choses et du monde, permettant de les représenter ainsi que d'accéder à leurs propriétés. Si les concepts désignent les représentations mentales de classes d'éléments, alors que les catégories désignent les classes elles-mêmes (Murphy, 2004) et si certains auteurs, comme Barsalou (1991), distinguent la fonction principalement inférentielle de la conceptualisation par rapport à celle organisationnelle de la catégorisation ; nous utiliserons ces termes de manière interchangeable dans la suite de ce texte.

Les questions de la nature et de la création des représentations mentales qui nous permettent d'identifier les objets comme appartenant à une certaine catégorie, ont été l'enjeu de plusieurs théorisations successives. Nous introduirons les quatre grandes familles, afin de présenter des caractéristiques importantes des catégories.

#### 1. THÉORIE CLASSIQUE DES CONCEPTS

Les théories fondées sur des règles furent les premiers modèles proposés (Hull, 1920 ; Bruner et al, 1956) et sont associées à ce qu'on appelle la conception classique des concepts (Smith & Medin, 1981), dans laquelle l'appartenance à une catégorie est déterminée par un ensemble de conditions nécessaires et suffisantes qui constituent la définition de la catégorie. La recherche s'accorde aujourd'hui sur ses insuffisances, dont les principales sont les suivantes :

Tout d'abord, il existe des catégories pour lesquelles il très difficile voire impossible de donner une définition en termes de conditions nécessaires et

suffisantes. Wittgenstein (1953) a ainsi donné l'exemple de la catégorie des « jeux » incluant aussi bien le « jeu d'échecs » que le « rugby ». Une autre limitation de l'approche classique est qu'elle exclut, en raison de la nature nécessaire et suffisante des propriétés définissant une catégorie toute distinction de degré d'appartenance de ses membres. Or de nombreux travaux empiriques, dont ceux de Rosh (1975, 1978 ; Rosch et Mervis, 1975), ont montré qu'une catégorie a en général une structure hétérogène et graduée allant des membres centraux aux atypiques et que plusieurs indicateurs permettent de distinguer les membres d'une catégorie selon leur typicalité.

Par ailleurs, l'approche classique suppose qu'à la question : "Tel objet appartient-il à telle catégorie ?», il doit toujours être possible de répondre par "oui" ou par "non" et exclut par conséquent toute ambiguïté d'appartenance à une catégorie. Ce qui a été contesté par les résultats empiriques de Hampton (1979), et de McCloskey et Glucksberg (1978). Une catégorie a des frontières floues (Sander, 2017).

Enfin, l'approche classique ne permet pas d'expliquer la non transitivité de certaines jugements catégoriels, comme ceux catégorisant "un siège de voiture comme un siège", "un siège comme un meuble", mais pourtant ne catégorisant pas "un siège de voiture comme un meuble" (Murphy, 2004).

## 2. THEORIE DES PROTOTYPES & THEORIE DES EXEMPLAIRES

La théorie des prototypes a été développée par Eleanor Rosch et ses collaborateurs (Rosch, 1975 ; 1978 ; Rosch & Mervis, 1975) pour rendre compte du phénomène de typicalité qu'ils ont observé lors de leurs expériences. Dans cette théorie, le jugement de l'appartenance à une catégorie s'effectue sur la base du degré de similarité avec le prototype, le "meilleur exemplaire" de la catégorie. Tous les membres d'une catégorie sont décrits par une liste de propriétés et le prototype correspond à la représentation des propriétés les plus fréquentes des membres de cette catégorie. Ainsi lorsque la catégorie « meubles » est considérée par exemple, l'élément « chaise » est plus fréquemment cité que celui de « tabouret ».

Les principales limites de cette théorie sont d'une part l'insuffisance de la représentation sous forme d'une liste de traits, qui ne rend compte que d'une partie seulement de l'information présente dans un concept; et d'autre part le fait que la représentativité et l'appartenance conceptuelle du membre de la catégorie ne soient pas distinguées et donc par conséquent que l'appartenance du membre devienne une question de degré (Murphy & Medin, 1985).

La théorie des exemplaires constitue une alternative à la théorie des prototypes (Hampton, 1993). L'idée que les sujets construisent une représentation abstraite qui incarne le concept est rejetée, cette théorie postule en revanche que sont stockés en mémoire tous les exemplaires rencontrés par le passé de cette catégorie et que leurs propriétés ont été encodées. Différents modèles (Medin et Schaffer, 1978 ; Brooks, 1978 ; Nosofsky, 1988) rendent compte de la catégorisation au moyen d'exemplaires, mais tous reposent néanmoins sur une comparaison des propriétés de l'élément à catégoriser avec celles des exemplaires encodés en mémoire. La typicalité d'un membre est définie en fonction de la quantité d'exemplaires avec lequel ils partagent une forte similarité. Ainsi un « rouge-gorge » par exemple est rapidement catégorisé comme appartenant à la catégorie « oiseau », car ressemblant en mémoire à de nombreux exemples de la catégorie « oiseau ».

Ces deux théories, dont les catégorisations sont fonction d'un degré de similitude soit avec le prototype soit avec des exemplaires mémorisés, souffrent néanmoins d'un défaut de contraintes sur ce qui est choisi comme propriétés pertinentes (Murphy & Medin, 1985), et ne permettent pas de rendre compte du phénomène de dépendance des catégories avec le contexte. En effet Barsalou (1982) a montré que les propriétés activées d'une catégorie dépendent du contexte : « le fait qu'un ballon de basket soit rond est activé dans tous les contextes, mais qu'il flotte l'est que dans des contextes spécifiques, par exemple des ballons dans un bateau » (Sander, 2016). La typicalité des exemplaires d'une catégorie dépend également du contexte. Barsalou (1983) s'est de plus concentrée sur l'étude des catégories “ad hoc” qui ne préexistent pas en mémoire à long terme et sont créées pour les besoins d'une tâche, telles que par exemple : « choses à mettre dans sa valise », « courses faire avant de partir en vacances ». L'existence de ces catégories ad hoc reflète la place essentielle de la

catégorisation dans la pensée humaine, l'absence d'une catégorie adéquate pour résoudre une tâche entraîne la création de cette catégorie. La construction de telles catégories est particulièrement importante lors la compréhension de situations ou de textes, en particulier dans le cadre de l'appréciation critique du monde qui nous entoure (Halpern, 2014).

### 3. THEORIE DES CONNAISSANCES

Cet exemple de catégorisation ad hoc permet de mettre en évidence les limitations de théories conceptuelles fondées sur le partage de traits. En effet, ce qui permet parfois de réunir les objets dans ces catégories ad hoc, peut tout à fait ne pas apparaître dans la description caractéristique de l'objet. Ainsi « un collier » et « un disque compact musical », jugés au départ comme appartenant à des catégories distinctes par les participants des expériences menées par (Barsalou, 1983) et citées par Sander (2016), peuvent être finalement associés comme membre de la catégorie commune « cadeaux possibles pour l'anniversaire d'une adolescente urbaine».

Une nouvelle approche, fondée sur les connaissances (Murphy, 2004) a été développée, afin de mieux rendre compte des liens profonds qui unissent les exemplaires d'une même catégorie, c'est-à-dire de la cohérence conceptuelle ainsi que du contexte dans lequel ils s'inscrivent. Ces modèles de catégorisation se fondent principalement sur l'explication. Ils permettent d'intégrer le fait que la cohérence de structure conceptuelle n'est pas toujours fondée que sur la similitude, ni sur le partage de propriétés comme dans les théories précédentes, mais également sur l'explication des relations entre les propriétés et leurs opérations possibles. Le concept qui fournit la meilleure explication des propriétés d'un élément à catégoriser représente la catégorie à laquelle cet élément appartient. Winsniecki et Medin (1994) offrent ainsi de nombreux exemples de catégorisation qui satisfont ce modèle : comme celui d'un convive sautant dans une piscine habillé sous l'effet de l'alcool, catégorisé dans la catégorie des “ personnes ivres”. La flexibilité de cette approche est rendue possible par la prise en compte des connaissances préalables et des croyances que les sujets ont sur le monde. Les représentations des catégories ressemblent davantage à des théories.

Ainsi le concept, d'abord conçu comme isolé dans une organisation hiérarchique et défini par des caractéristiques propres, est inséré dans des structures contextualisées, puis dans l'ensemble du domaine des connaissances et croyances (Murphy, 2004). Toutes ces connaissances relationnelles sont liées et s'influencent mutuellement. L'élaboration des concepts est ainsi étroitement liée aux théories dites naïves, implicites qu'un sujet possède au sujet d'un domaine particulier (Keil, 1989 ; Murphy et Medin, 1985).

Les concepts, qui se développent tout au long de la vie d'une personne, s'ajustent selon ses expériences, appréhendées elles-mêmes en fonction des concepts dans leur état antérieur. Le développement conceptuel repose ainsi sur des mécanismes de comparaison mentale, d'analogie, qui permettent d'aborder la nouveauté en mobilisant ses expériences passées (Sander & Hofstadter, 2013).

## B. COMMENT CATEGORISONS-NOUS LES SITUATIONS ?

### 1. MODELES DE COMPREHENSION DE TEXTE

Comprendre un texte nécessite d'aller au-delà de ce qui est explicitement exprimé et de saisir la signification de ce qui est véhiculé par le texte (Gough & Tunmer, 1986). La compréhension d'un texte peut se définir comme la construction d'une succession de représentations mentales sur plusieurs niveaux. Le modèle de Kirsch et Van Dijk (1973), qui a servi de base aux modèles de compréhension de texte, considère ainsi deux niveaux de représentation : le premier, qui est appelé base texte, correspond à l'analyse propositionnelle de chacune des phrases du texte ; le deuxième, la macrostructure, qui est contrôlée par l'intention de lecture du sujet, organise l'ensemble de ces propositions en une structure cohérente pour en faire ressortir le sens. Sur cette base, de nombreux modèles de compréhensions de textes ont par la suite été développés. Nous en présenterons les deux grandes familles, qui jouent un rôle important dans la catégorisation de situations.

### a) MODELES DES SCHEMAS

Dans leur modèle Kintsch et Van Dijk font intervenir la notion schéma, qui est considéré comme la représentation formelle de l'intention de lecture. Le concept de schéma, qui a été utilisé par Rynnart et Ortony (1977) ainsi que par Minsk (1975) avec la notion de cadre et par Schank et Abelson (1977) avec celle de script, désigne l'ensemble des connaissances acquises et stockées en mémoire à long terme sous forme de structure générale, qui est utilisé dans les processus de compréhension, de mémorisation ou de production d'inférences. Dans un modèle de schéma, comprendre un texte consiste à construire une représentation de la situation en activant au cours de la lecture un schéma de connaissances en mémoire et en y instanciant certains éléments du texte (Richard, 1995). Le schéma va permettre de rétablir les informations manquantes et les intégrer dans un réseau de relations non nécessairement explicitées par le texte. Ce modèle de compréhension de texte a été élargi par Kintsch et Greemo (1985) dans le contexte de la résolution de problèmes arithmétiques. En effet, la compréhension de textes ne se limite pas au domaine littéraire, elle joue un rôle dans la résolution de problèmes mathématiques à énoncés verbaux.

Toutefois cette approche des schémas a été beaucoup controversée car ne permettant pas de rendre compte de l'interprétation du sujet ainsi que des effets liés au contexte.

### b) MODELES DE SITUATION

Afin d'y remédier, Reusser (1990) introduit le modèle de situation dans le processus de compréhension en tant que représentation intermédiaire, construite à partir des connaissances générales du sujet et des éléments contenus dans le texte. Les agents, les évènements et les relations entre les entités décrites dans le texte y sont spécifiés. Kintsch (1988) intègre le modèle de situation dans la deuxième version de son modèle de compréhension, comme étant composé de la base texte et des inférences sur la situation faites par le lecteur à partir de ses connaissances générales et expériences vécues.

L'avantage important de cette approche est qu'elle permet la prise en compte des effets de contexte ainsi que des aspects interprétatifs observés lors de la compréhension de texte. Les travaux plus récents (Kintsch, 1992, McKoon & Ratcliff, 1992) ont mis évidence que les sujets élaborent des représentations de niveau d'abstraction élevé, inférées au cours de la lecture.

La compréhension d'un texte, que cela soit par l'activation d'un schéma stocké en mémoire ou par la construction d'un modèle de situation intégrant des éléments activés par des connaissances préalables, repose donc sur des mécanismes d'analogie.

## 2. L'ANALOGIE

L'analogie permet en s'appuyant sur des similitudes perçues, d'aborder l'inconnu à partir de ce que l'on connaît (Sander, 2000). En effet, pour apprêhender tout nouveau concept ou nouvelle situation, l'être humain fait appel à des éléments source connus sur lesquelles il va s'appuyer pour donner du sens au concept ou à la situation cible. L'analogie joue donc un rôle essentiel dans l'acquisition des connaissances. De nombreuses études ont porté sur l'analogie au cours de ces quatre dernières décennies, souvent basées sur la résolution de problèmes. Les principales théories élaborées sont les suivantes :

Selon Gentner (1983, 1989), l'analogie repose avant tout sur une similitude des structures relationnelles de la source et de la cible. Le transfert porte sur le système de relations entre des entités, indépendamment des propriétés spécifiques des entités dans chacune des situations. Holyoak et ses collaborateurs (1985), quant à eux reprochent à cette théorie d'ignorer le contexte de la situation ainsi que la réalisation des buts du sujet. Ils apportent une alternative pragmatique basée sur la théorie des schémas, dans laquelle les deux analogues forment par leurs éléments communs un schéma plus général et abstrait. Au fil du temps, les deux théories se rapprochent et Holyoak va intégrer l'approche syntaxique de Gentner dans son modèle.

De nombreuses études se sont concentrées sur les différentes phases du raisonnement par analogie, dont le nombre d'étapes varie selon les auteurs. La

plupart s'accordent néanmoins sur le passage par les phases d'évocation, d'appariement et d'utilisation décrites par Ripoll (1992). L'essentiel des travaux se sont focalisés sur les étapes d'évocation et d'appariement et sur le rôle des traits de surface et de structure dans ces processus.

Les traits de surface représentent les caractéristiques de la situation non pertinentes pour l'analogie : les informations relatives au texte, le thème général, les évènements, les agents ou objets, les actions, etc... alors que la structure profonde correspond au squelette conceptuel, au scenario abstrait de l'histoire décrite dans le texte. Si la plupart des auteurs s'accordent sur le fait que le transfert analogique est rarement spontané lors d'isomorphismes structurels, que des incitations explicites sont souvent nécessaires (Gick & Holyoak, 1980 ; 1983 ; Holyoak et Koh, 1987) et sur l'importance des traits de structure dans le processus d'appariement, il n'y a pas consensus en ce qui concerne la phase d'évocation (Blanchette & Dunbar, 2000 ; Dunbar et Blanchette, 2001 ; Hofstadter & Sander, 2013 ; Raynal, Clément et Sander, 2017), bien que certains résultats empiriques mettent en évidence la grande influence de l'habillage de surface (Gick & Holyoak, 1983 ; Gentner, Ratterman & Forbus, 1993 ; Trench & Minervino, 2015).

La difficulté du transfert lors d'isomorphismes structurels peut s'expliquer par une incapacité à encoder la situation à un niveau suffisamment abstrait pour en saisir les similarités (Gentner et al, 2009 ; Medin & Ross, 1989 ; Hofstadter & Sander, 2013). Des travaux montrent que les représentations construites par le sujet dépendent à la fois de ses connaissances préalables et du niveau d'abstraction des similarités. Les travaux de Novick (1988) ont ainsi mis en évidence que cet effet de surface est moins prononcé lorsque l'individu possède l'expertise suffisante pour extraire une structure abstraite.

Bassok (2001) quant à elle rejette cette dichotomie entre habillage de surface et structure profonde et considère que plusieurs conceptualisations expertes ou catégorisations pertinentes d'une même situation sont possibles. Elle propose une hypothèse interprétative, dans laquelle une structure induite est construite à partir des éléments de surface et des connaissances préalables. Cette structure induite, sur laquelle se base l'interprétation du sujet, peut se définir comme

l’abstraction relationnelle qui émerge des objets décrits dans le texte, un niveau intermédiaire entre les éléments de surface et le squelette conceptuel du texte. Cette contribution est essentielle car elle permet rendre compte des connaissances préalables du sujet quant aux concepts et relations décrites dans le textes et de comprendre comment une même histoire ou contenu peuvent donner lieu à des interprétations différentes suivant le contexte ou les connaissances du sujet.

Chaque être humain fabrique sans cesse de nouvelles catégories ad hoc lors de l’appréhension de situation (Barsalou, 1983) et par conséquent de leur encodage. Cet encodage peut se faire à plusieurs niveaux d’abstraction, et la catégorie sélectionnée détermine la perspective selon laquelle la situation va être appréhendée. Sander (2000) et Ripoll (2001) remettent en question dans les théories précédentes le rôle central de l’étape de mise en correspondance, qui présuppose que la représentation de la cible soit construite avant la réalisation de l’analogie et que la représentation initiale de la cible soit élaboré à un niveau d’abstraction équivalente à celle de la source. Il est en effet légitime de considérer que la source fournit des informations sur la catégorie potentielle à laquelle peut être associée la cible. En outre, selon Sander (2000), l’analogie est impliquée dès l’encodage de la cible dans les termes de la source. En d’autres termes, le sujet encode des situations selon les dimensions, qui ultérieurement détermineront quelles autres situations conduiront à les évoquer (Sander & Hofstadter, 2013). De plus, avec une approche convergente avec celle de la structure induite de Bassok, Sander et Hofstadter (2013) mettent en exergue le fait que pour que l’évocation soit réussie il n’est pas nécessaire que la source ait été encodée au niveau d’abstraction le plus élevé et/ou que la catégorie abstraite dont les analogues constituent des exemplaires préexiste en mémoire, il suffit que de nombreux aspects importants et abstraits aient été encodés et favorisent l’analogie.

## C. COMMENT DEVELOPPER NOTRE ESPRIT CRITIQUE ?

L'opération mentale de catégorisation fournit un éclairage particulier sur une situation, permet d'en construire une représentation et d'inférer des hypothèses. Réciproquement, l'immersion dans des situations va guider la construction de nouvelles catégories et la transformation de celles existantes. La catégorisation représente donc un processus fondamental de prise d'information sur le monde et de compréhension (Hofstadter & Sander, 2013).

### 1. EXPERTISE

La spécificité des experts réside dans la manière dont ils se représentent les informations dans leur domaine d'expertise et les catégorisent. Les travaux princeps de Chi, Feltovitch et Glaser (1981) ont montré que ce qui distingue des experts en physique de novices est la façon dont ils élaborent leurs représentations en se basant sur les éléments de structure versus uniquement sur ceux de surface. Alors que les novices regroupent des problèmes faisant intervenir les mêmes objets, comme des poulies ou des plans inclinés, les experts catégorisent les problèmes selon le principe physique mis en jeu, comme la troisième loi de Newton par exemple, tout en tenant compte des traits de surface. Différents travaux confortent ces résultats et mettent en évidence le fait que le niveau d'expertise influence la nature de la catégorisation, qu'il s'agisse des critères de classification (Boster & Johnson, 1989), de la nature graduée des catégories et la typicalité de ses membres (Rosch, 1973 ; Medin et al, 1997) ; mais aussi leur degrés d'abstraction (Rosch et al ; 1976) ainsi que leur hiérarchisation (Akin, 1992 ; Chatard-Pannetier, 2002). L'expert n'a pas seulement multiplié ses connaissances et catégories, il les a organisées de manière à pouvoir passer de l'une à l'autre au besoin et de trouver celle au niveau d'abstraction le plus pertinent suivant le contexte. Ainsi selon Sander (2013), « l'expert est celui qui sait abstraire l'habillage pour percevoir l'essence ».

L'expertise n'est pas spécifique à des sujets techniques ou des secteurs en particulier, mais fait référence ici à toute connaissance approfondie et savoir-faire que l'individu peut avoir développée dans un domaine, y compris son environnement quotidien. Le terme expert (du latin *expertus*), adjectif au XIII<sup>e</sup> siècle, désignait la personne, qui par l'expérience et la pratique avaient acquis une grande habileté. Dans le langage courant, nous le distinguons du savant ou du spécialiste en ce sens que l'expert possède des connaissances mais est amené à résoudre un problème donné. L'expert est supposé avoir à la fois les connaissances et le savoir-faire approprié (Dubois et al , 2015).

Les études portant sur l'expertise (Adelson, 1981 ; Chase et Simon, 1973 ; Chi, Glaser et Farr, 1988 ; Ericsson and Simon, 1993, Robinson and Hayes, 1978, Wineburg, 1991, Glaser, 1992), quel que soit le domaine concerné, permettent de mieux comprendre le fonctionnement cognitif des individus, d'identifier les caractéristiques propres aux experts et de la régularité du « parcours » pour le devenir. L'apprentissage vise de nos jours à faire apprendre aux sujets non seulement des connaissances mais des connexions entre concepts, des critères de classification semblables à ceux des experts dans le domaine (AAAS “Vision and Change” report, 2011).

## 2. ESPRIT CRITIQUE

Les experts semblent non pas se distinguer par des différences dans leurs fonctions cognitives mais par des capacités accrues d'abstraction et de flexibilité cognitive dans leur domaine d'expertise, associées avec la reconnaissance des catégories les plus pertinentes. Ces notions s'intègrent tout à fait dans le cadre de l'appréciation critique du monde qui nous entoure.

Si le développement de la pensée critique a toujours constitué une mission de l'école, il représente aujourd'hui un objectif primordial mis en avant par le gouvernement, comme le dit une page spécialisée d'Educsol « le développement de l'esprit critique est au centre de la mission assignée au système éducatif français » et un grand intérêt est porté aux méthodes et outils pour enseigner l'esprit critique au sein des curriculum scolaires sont de plus en plus populaires.

Selon la fondation américaine de l'esprit critique, la pensée critique est « un processus intellectuellement discipliné, qui consiste à conceptualiser, appliquer, analyser, synthétiser ou à évaluer de façon active et habile l'information recueillie ou générée par suite de l'observation, de l'expérience, de la réflexion, du raisonnement et de s'en servir pour croire ou agir ». Il n'y a pas de réel consensus autour de la définition même de l'esprit critique et celle que nous venons de citer est essentiellement philosophique. Nous lui préfèrerons l'approche de Halpern (1998), qui définit l'esprit critique par référence aux cinq grands groupes de capacités cognitives que sont : la compréhension des énoncés verbaux, l'argumentation, le raisonnement hypothético-déductif, l'évaluation de l'incertitude et la résolution de problèmes. Dans le cadre de ce travail, nous nous focaliserons essentiellement sur la compréhension verbale et la résolution de problèmes ou l'analogie ainsi que la capacité de percevoir la structure profondes des situations sont fondamentales.

L'une des questions importantes qui est posée, est de savoir si l'esprit critique est une compétence transversale regroupant un ensemble de savoir-faire qui pourraient être transférés à n'importe quel contenu. Si certaines approches autonomes se basent sur l'enseignement de la pensée critique en tant que discipline en soi, d'autres soulignent la difficulté d'isoler et de transférer des compétences critiques indépendamment des domaines d'expertise. Les études empiriques montrent que ces compétences dépendent du contexte, dont les connaissances et le contenu sur lequel elles s'exercent (Willingham, 2009) ; mais qu'il est néanmoins possible de favoriser le transfert des apprentissages par la mise en place de stratégies spécifiques (Perkins & Salomon, 1988 ; Bransford et al, 2000).

Les sciences, qui se fondent sur la construction de connaissances bien fondées et solides appuyées sur des faits et hypothèses, ont un intérêt tout particulier pour l'esprit critique, qui sous-entend une prise de distance par rapport à un contenu portant à réflexion et la sélection du point de vue le plus pertinent parmi plusieurs possibles (Halpern 2003 ; 2010). Les méthodes et outils scientifiques permettent ainsi de combattre les heuristiques et biais spontanés dont nous sommes victimes et d'outiller la pensée critique au quotidien (Farina, Pasquinelli et Zimmerman, 2018). L'esprit critique va de pair

avec la pensée rationnelle, qui se distingue de l'intelligence telle que mesurée par le QI (Kahneman, 2012 ; Stanovitch, 2014 ; 2016).

### III. Problématique et hypothèses

---

La difficulté du transfert lors d'isomorphismes structurels peut s'expliquer par une incapacité à encoder la situation à un niveau suffisamment abstrait pour en saisir les similarités. Avec une perspective analogue aux travaux de Novick (1988), Raynal, Clément et Sander (2017) défendent l'hypothèse que l'évocation se fonde sur les similarités structurelles lorsque les connaissances du sujet lui permettent d'encoder les dimensions abstraites de la situation et de négliger les détails de surface. Pour le montrer, alors que la plupart des travaux de recherche se sont jusqu'à présent centrés sur des analogies faisant intervenir des domaines spécifiques dont les concepts ne sont pas connus des participants, Raynal et al (2017) font appel à des petits récits décrivant des scènes de vie quotidienne dont les concepts sont familiers pour les participants.

Les tâches classiquement utilisées pour étudier le transfert analogique sont :

- de résolution de problème puzzle ou mathématiques (Gick et Holyoak, 1983 ; Novick, 1988 ; Ross, 1987), dans laquelle un problème source et sa solution sont initialement présentés au sujet qui doit ensuite résoudre un problème cible analogue ;
- de rappel d'histoires (Gentner et al, 1993 ; Raynal, Clement et Sander, 2017), dans laquelle une liste de textes sources est initialement présentée au sujet qui doit par la suite essayer de s'en souvenir en lisant une liste de problèmes cibles présentant des similarités de surface et/ou de structure avec certains textes sources;
- de production d'analogies (Blanchette and Dunbar, 2000; Dunbar et Blanchette, 2001; Trench et Minervino, 2015), dans laquelle le sujet doit évoquer à partir d'une situation cible donnée des expériences analogues qu'il a vécues de par le passé.

D'autres auteurs se sont intéressés plus particulièrement aux représentations mentales ainsi qu'à l'organisation des connaissances des sujets par le biais de tâches de classification de problèmes scientifiques. Dans la lignée des travaux de Chi, Feltovitch et Glaser (1981), plusieurs études (Smith et al, 2013 ; 2017 ; Krieter, Julius, Tanner, Bush et Scott, 2016) ont été menées ces dernières années pour développer des outils pédagogiques permettant d'évaluer non pas l'acquisition de tel ou tel concept scientifique mais l'organisation conceptuelle des connaissances des étudiants dans certaines matières (notamment en chimie et en biologie). Des exercices de catégorisation utilisant des cartes pouvant être classées suivant soit des caractéristiques de surface, soit des concepts scientifiques plus abstraits ont ainsi été élaborés pour évaluer les différences d'expertise conceptuelle des étudiants dans ces matières.

S'inscrivant dans la lignée de ces études et des travaux de Raynal, Clément et Sander, notre étude s'articule autour de tâches de catégorisation et de sélection d'analogies impliquant des petits récits de la vie quotidienne. Nous nous sommes intéressés à des populations d'élèves de 6<sup>eme</sup> se distinguant par leurs performances scolaires et par le fait que ces élèves aient été diagnostiqués ou non comme étant précoces. En outre, notre objectif est de mettre en évidence chez les élèves, des caractéristiques communes/ distinctes en fonction de la manière dont ils se représentent l'information et la nature des critères de catégorisations utilisés dans des situations du quotidien dont ils disposent tous des concepts abstraits. En effet, chaque élève peut en théorie devenir expert de son quotidien, indépendamment de connaissances scolaires acquises. En utilisant ces stimuli nous souhaitons extraire des informations quant aux niveaux d'abstractions de leurs catégorisations ainsi qu'à l'organisation de leur répertoire conceptuel.

La première tâche de notre protocole expérimental repose sur une catégorisation de textes dans laquelle les énoncés peuvent être classés suivant soit des caractéristiques de surface, soit des schémas conceptuels plus abstraits, qui constituent leur structure. Ces textes ont été élaborés de manière à ce que deux catégorisations exclusives l'une de l'autre, fondée pour l'une sur les éléments de surface et pour l'autre sur les éléments de structure, permettent de regrouper/ différencier l'ensemble des situations. La deuxième tâche est une sélection d'analogies visant à évaluer le niveau maximal d'abstraction des similarités perçues par chaque élève. Cet exercice spécialement élaboré en complément de la première tâche permet de recueillir des informations relatives au cas où les similarités de structure profonde sont perçues mais non utilisées comme critères de catégorisation dans la tâche 1.

Nous avons ainsi formulé les hypothèses que d'une part, les élèves dont les performances scolaires sont les meilleures ont plus tendance à percevoir et à guider leur classification suivant des éléments de structure et moins sur des éléments de surface que les élèves moins performants ; et que d'autre part, cette capacité est d'autant plus développée par le groupe d'élèves à Haut Potentiel. Ces tâches pourraient ainsi contribuer à la constitution d'un outil pédagogique permettant de mettre en évidence et de quantifier des différences concernant le niveau d'abstraction des structures perçues par les élèves et les dimensions selon lesquelles ils encodent les situations.

## IV. Méthodologie

---

### A. PARTICIPANTS

Cette recherche a été réalisée dans 16 classes de 6ème de cinq collèges situés à Paris et à Orléans. Trois établissements ont été sélectionnés parmi les collèges du réseau de la Map en fonction de l'intérêt des professeurs et chefs d'établissements. Les deux autres ont été choisis car scolarisant de nombreux enfants diagnostiqués à Haut Potentiel Intellectuel, « HP », dans l'ensemble de ses classes de 6ème. Notre sous-groupe d'élèves HP est constitué des 42 élèves de 6ème de ces deux établissement ayant été diagnostiqués « Haut Potentiel » par une psychologue agréée, à la suite de l'obtention d'un score de  $QI > 130$  aux tests psychométriques WISC IV ou WPP III (en excluant les profils trop hétérogènes) et dont le bilan a été communiqué à la direction du collège.

**Stratégie d'échantillonnage :** Étant donné que l'ensemble des enfants diagnostiqués HP proviennent uniquement de deux établissements et que ces établissements présentent des distinctions importantes (établissement privé, éducation bilingue, nombres de classes ...) par rapport aux trois autres, nous avons décidé pour la cohérence de cette étude de stratifier notre étude en deux parties et notre échantillon total en deux sous-échantillons : l'échantillon 1 correspondant aux 201 élèves du Lycée 1 et 2, et l'échantillon 2 aux 205 élèves des lycées 3,4,5.

Toutes les analyses comparatives des sous-groupes HP et Standard menées afin de répondre à nos questions de recherche, se sont basées uniquement sur les données de l'échantillon 1, alors que celles étudiant la corrélation avec le niveau de performance scolaire ont utilisées celles de l'échantillon 2. Pour compléter notre étude, nous avons de plus tenté de caractériser pour l'échantillon 1 la relation entre nos variables dépendantes et chacun des deux facteurs Niveau Scolaire et Diagnostic l'influencant, en tenant compte de leur interaction.

Dans chacun des collèges, l'ensemble des classes de 6ème ont participé à l'expérience, à l'exception des absents le jour de la passation ainsi que des élèves dont les parents ont refusé qu'ils prennent part à cette étude. 413 élèves ont ainsi été sollicités ; de ceux-ci, cinq n'ont pu y prendre part pour les raisons évoquées et deux n'ont pas remis leurs réponses. Notre échantillon total (resp. l'échantillon HP) se compose donc de 406 élèves de 6ème (resp. 42 élèves), répartie entre 210 garçons et 196 filles (resp. 28 garçons et 14 filles), de toutes catégories socio-culturelles confondues.

Nous avons recueilli auprès des professeurs des données concernant les élèves ayant passé l'expérience. Tout en respectant l'anonymat des élèves via une identification par numéro, nous avons demandé aux professeurs de remplir un questionnaire indiquant les âge, sexe, langue maternelle des élèves, s'ils avaient un handicap pour la lecture et s'ils avaient été diagnostiqués HP. Nous leur avons également demandé de répartir leurs élèves suivant trois niveaux A, B et C en fonction des résultats du conseil de classe afin d'obtenir un classement aussi objectif que possible (ne dépendant pas du jugement d'un seul professeur), bien que n'étant pas construit selon les mêmes critères suivant les lycées. Le tableau 4.A.1 représente les répartitions des deux échantillons suivant les sous-groupes qui ont été étudiés.

Eleves	Echantillon 1						Echantillon 2				
	HP		Performances Scolaires			Total	Performances Scolaires			Total	
	Non-Diagnost	Diagnostic	A	B	C		A	B	C		
Sexe	M	74	28	47	43	12	102	37	38	33	108
	F	85	14	46	46	7	99	30	38	29	97
	Total	159	42	93	89	19	201	67	76	62	205
Age	10	2	6	4	4	0	8	2	0	1	3
	11	96	24	49	56	15	120	37	33	29	99
	12	61	12	40	29	4	73	28	40	24	92
	13	0	0	0	0	0	0	0	2	8	10
	14	0	0	0	0	0	0	0	1	0	1
	Total	159	42	93	89	19	201	67	76	62	205

Tableau 4.A.1: Repartition des élèves par sexe et par age au sein des sous-populations d'élèves

## B. MATERIEL

L'ensemble des textes utilisés dans cette étude sont des petits récits de la vie quotidienne présentant des similitudes de surface ou de structure profonde. Spécialement conçus pour les deux tâches de notre protocole expérimental, leur lecture et compréhension sont accessibles pour des enfants de 6ème. Le vocabulaire employé, la longueur des textes ainsi que le format de présentation utilisé ont été contrôlés en amont avec l'ensemble des professeurs pour s'en assurer. Les thèmes et scénarios illustrés par ces textes ont été pensés en les choisissant proches du vécu des enfants de cet âge : toutes les situations décrites sont liées à une activité scolaire ou sportive ou à des loisirs entre amis.

Le protocole expérimental conçu pour répondre à nos questions de recherche comprend deux tâches distinctes. Chacune d'entre elles permet de recueillir des données sur les représentations mentales des élèves et leur capacité à percevoir des similarités de structure correspondant à différents niveaux d'abstraction.

La première tâche est une tâche de catégorisation. Un paquet de 9 textes écrits sur des cartes plastifiées détachées était distribué à chaque élève ainsi qu'une petite pochette contenant élastiques et post-it (voir photos en Annexes). La consigne était de regrouper les textes exprimant la même idée. Il était précisé aux élèves qu'ils pouvaient réaliser autant de catégories qu'ils le souhaitaient, chaque catégorie pouvant regrouper de 1 à 9 cartes mais chaque carte ne pouvant appartenir qu'à une seule catégorie. Dans chaque jeu de stimuli, trois structures superficielles et trois structures profondes sont représentées, chaque texte incorporant une des trois structures superficielles et une des trois structures profondes. Il y a par conséquent trois fois trois textes incorporant les mêmes traits de surface et trois fois trois textes reflétant la même structure profonde, qui représentent les deux catégorisations potentielles mutuellement exclusives que nous nommerons (SS) et (SP). L'ensemble des textes ont été élaboré en tentant d'éviter toute autre similarité entre les différents textes, afin qu'aucune autre catégorisation saillante n'entre en compétition avec ces deux catégorisations hypothétiques.

Afin de maximiser leur implication dans la tâche et éviter toute catégorisation aléatoire, la consigne demandait de plus aux élèves de donner un titre à chaque catégorie qu'ils avaient formée : un mot, une phrase qui expliquait pourquoi ils avaient regroupé les différents textes ensemble. Les élèves devaient rassembler chaque groupe de cartes à l'aide des élastiques puis inscrire le titre correspondant sur un post-it collé dessus.

La seconde tâche est une tâche de détection d'analogie entre situations. Après avoir lu un texte source chaque participant devait identifier dans une liste de quatre textes (cibles), ceux où l'on retrouvait la même idée que dans le texte source. L'ensemble des textes était présenté sur une feuille de papier A4 et l'instruction donnée était de cocher la case du texte cible correspondante (voir Tâche 1.A p. 31). La liste des textes cibles a été construite de manière à comporter trois textes de degrés croissant d'abstraction en termes de relations de similitudes (niveau 1, niveau 2, niveau 3), ainsi qu'un texte « intrus » ne présentant aucune similitude (ni de surface ni profonde).

Le positionnement des quatre textes cibles a été randomisé tout en satisfaisant certaines contraintes afin d'éviter certains biais dans les réponses : le texte intrus n'était pas positionné en bordure et les histoires n'étaient pas présentées dans un ordre crescendo ou decrescendo en degrés d'abstraction dans leurs relations de similitudes. Afin de pallier aux difficultés associées à la subjectivité du classement des niveaux d'abstraction des similarités (problème rencontré lors de la phase d'élaboration et du premier test de notre matériel), les énoncés ont été construits de manière à présenter des similarités hiérarchisées et inclusives et donc permettre un classement objectif. L'analogue de Niveau 1 comporte des similitudes de surface, intermédiaire et profonde avec le texte source, l'analogue de Niveau 2 ne comporte que celles intermédiaire et profonde et finalement l'analogue de Niveau 3 que la profonde (voir Tâche 2.A p. 34).

Pour chaque tâche, plusieurs versions des stimuli ont été construites (deux pour la tâche 1, trois pour la tâche 2) afin d'éviter tout biais lié spécifiquement aux énoncés et de vérifier la cohérence interne des réponses intra sujets. L'ensemble des versions créées pour les deux tâches sont présentées en annexes

mais sont détaillées dans les pages qui suivent, une version des stimuli pour chaque tache à titre d'exemple.

Dans la version A de la tâche 1, les trois structures de surface et les trois structures profondes sont les suivantes (voir p. 32-33) :

- Structures de surface : le foot, les jeux vidéo et les devoirs.
- Structures profondes : l'union fait la force, reculer pour mieux sauter, réussir par chance alors que tout semble perdu.

Dans la version A de la tâche 2 (voir p. 34-35), la structure profonde commune aux trois analogues est “le retour au point de départ malgré des efforts pour avancer”. L'intrus ne présente aucune similarité avec le texte source à quelque niveau d'abstraction que cela soit. L'analogie de niveau 1 partage en plus de la structure profonde, les similarités de surface que constituent le thème du retour chez soi le plus rapidement possible. Les analogues de niveau 2 et de niveau 3 quant à eux se distinguent par la nature de plus en plus abstraite et éloignée de celui du texte source du monde dans lequel se décline le schéma conceptuel.

Les hypothèses intrinsèques à la construction de cette tâche, sont d'une part que l'ensemble des participants est en mesure d'identifier l'analogie de niveau 1 (incorporant à la fois des similarités de surface, de structure intermédiaire et profonde avec le texte source) et d'autre part qu'en raison de la nature emboîtée des analogies de degré d'abstraction décroissant, tout élève qui identifie l'analogie de niveau n, devrait aussi avoir identifié ceux de niveau n-1 et n-2. Ainsi de manière générale sur l'ensemble total des élèves, nous anticipons des proportions d'analogie perçus strictement décroissantes suivant les niveaux d'abstraction.

Deux groupes « test » de 10 enfants âgés entre 10 et 13 ans ainsi que deux groupes « plafond » de 10 adultes de l'ENS ont été constitués (pour passage individuel de la tâche, à un mois d'intervalle pour chaque paires de groupes) lors de la phase d'élaboration des stimuli afin de vérifier la non-ambiguïté du matériel et la cohérence de nos hypothèses.

### TACHE 1.A:

Lisez attentivement l'ensemble des textes distribués. Regroupez ensemble ceux qui expriment la même idée générale, puis trouvez un titre pour chaque groupe.

Manon est bloquée dans son jeu vidéo. Découragée, elle arrête et décide de regarder la télé.

En zappant, elle tombe sur une émission sur les jeux vidéo qui explique la solution de son problème.

Maxime a un contrôle mais n'a pas assez travaillé. Il n'a appris qu'une seule leçon sur les six qu'il fallait réviser. C'est précisément sur cette leçon que le professeur l'interroge. Il réussit son contrôle.

Avant d'affronter le méchant de la fin de son jeu vidéo, Noémie préfère retourner dans les niveaux précédents pour gagner des vies supplémentaires.

Emma a terminé son devoir et pourrait le rendre tout de suite ainsi, mais elle a fait beaucoup de ratures en rédigeant. Elle choisit de le recopier proprement sur une nouvelle feuille pour avoir une meilleure note.

Alors qu'il est devant les buts, Clément rate son tir. Le ballon part sur le côté mais rebondit sur le ventre d'un joueur de l'autre équipe et rentre dans les buts.

Lucas adore jouer aux jeux vidéo. Lorsque ses copains sont avec lui pour l'encourager, Lucas joue toujours mieux que lorsqu'il est seul.

Seuls devant leur copie, les élèves n'ont pas réussi à trouver la solution de l'exercice. Mais lorsque la maîtresse les a fait travailler en petits groupes, toujours sans leur donner aucune indication, plus de la moitié des groupes a résolu le problème.

Thomas n'arrive pas à dribler les 2 défenseurs en face de lui. Il passe le ballon à son coéquipier, qui lui redonne le ballon une fois qu'il s'est démarqué. Thomas tire et marque le but.

Manon est bloquée dans son jeu vidéo. Découragée, elle arrête et décide de regarder la télé. En zappant, elle tombe sur une émission sur les jeux vidéo qui explique la solution du problème.

Avant d'affronter le méchant de la fin de son jeu vidéo, Noémie préfère retourner dans les niveaux **LES JEUX** gagner des vies supplémentaires. **VIDÉO**

Maxime a un contrôle mais n'a pas assez travaillé. Il n'a appris qu'une seule leçon sur les six qu'il fallait réviser. C'est précisément sur cette leçon que le professeur l'interroge. Il réussit son contrôle.

Emma a terminé son devoir et pourrait le rendre tout de suite ainsi, mais elle a fait beaucoup de ratures en rédigeant. Elle choisit de le recopier proprement sur une **TOUS DEVOIRS** une meilleure note.

Alors qu'il est devant les buts, Clément rate son tir. Le ballon part sur le côté mais rebondit sur le ventre d'un joueur de l'autre équipe et rentre dans les buts.

Inquiété par les joueurs adverses, Pierre préfère faire une passe arrière à son gardien, plutôt que d'avancer. Celui-ci dégage le ballon dans l'autre moitié du terrain. Pierre **FOOT** le terrain en courant. Arrivé près des buts adverses, il récupère le ballon pour tirer.

Thomas n'arrive pas à dribler les 2 défenseurs en face de lui. Il passe le ballon à son coéquipier, qui lui redonne le ballon une fois qu'il s'est démarqué. Thomas tire et marque le but.

Seuls devant leur copie, les élèves n'ont pas réussi à trouver la solution de l'exercice. Mais lorsque la maîtresse les a fait travailler en petits groupes, toujours sans leur donner aucune indication, plus de la moitié des groupes a résolu le problème.

Manon est bloquée dans son jeu vidéo. Découragée, elle arrête et décide de regarder la télé. En zappant, elle tombe sur un émission sur les jeux vidéo qui explique la solution du problème.

Maxime a un contrôle mais n'a pas assez travaillé. Il n'a appris qu'une révision. C'est précisément sur cette leçon que le professeur l'interroge. Il réussit son contrôle.

Alors qu'il est devant les buts, Clément rate son tir. Le ballon part sur le côté. Mais il rebondit sur le ventre de Maxime. C'est précisément sur cette leçon que le professeur l'interroge. Il réussit son contrôle.

Avant d'affronter le méchant de la fin de son jeu vidéo, Noémie préfère retourner dans les niveaux précédents pour gagner REÇUTER POUR MIEUX SAUTER sur une nouvelle feuille pour avoir une meilleure note.

Inquiétée par les joueurs adverses, Pierre préfère faire une passe arrière à son gardien, plutôt que d'avancer. Celui-ci récupère le ballon dans l'autre moitié du terrain. Pierre traverse tout le terrain en courant. Arrivé près des buts adverses, il récupère le ballon pour tirer.

Lucas adore jouer aux jeux vidéo. Lorsque ses copains sont avec lui pour l'encourager, Lucas joue toujours mieux que lorsqu'il est seul.

Seuls devant leur copie, les élèves n'ont pas réussi à trouver la solution de l'exercice. Mais lorsque la maîtresse L'UNION FAIT LA FORCE, toujouors sans leur donner aucune indication, plus de la moitié des groupes

Thomas n'arrive pas à dribler les 2 défenseurs en face de lui. Il passe le ballon à son coéquipier, qui lui redonne le ballon une fois qu'il s'est démarqué. Thomas tire et marque le but.

à résolu le problème.

## TACHE B.1

Lisez attentivement le texte suivant.

En sortant du restaurant, les parents de Théo veulent prendre un raccourci pour rentrer chez eux. Après avoir marché pendant plus de 15 minutes, les voilà de nouveau devant l'entrée du restaurant !

Puis lisez séparément chacun des textes ci-dessous. A chaque fois qu'un texte exprime la même idée que le texte que du haut, cochez la case correspondante. Vous pouvez cochez plusieurs cases ou aucune.

Léa est très pressée et décide de rentrer chez elle en bus plutôt qu'en marchant. Elle prend le bus à la station devant la gare. A peine assise, elle s'endort et manque son arrêt. Lorsqu'elle rouvre les yeux, le bus est revenu devant la gare.

Yanis a rendez-vous à la piscine avec ses amis. Sa mère le conduit en voiture mais ils sont coincés dans les bouchons. Il descend de la voiture et finit le trajet en courant. Il arrive tout de même à l'heure à son rendez-vous.

Sacha a un problème de physique à résoudre, il ne trouve pas la solution. Depuis plusieurs heures il travaille à son bureau, mais il n'a rien écrit sur sa feuille. De nombreuses idées lui viennent à l'esprit mais aucune ne fonctionne.

Chloé souhaite se changer pour être plus élégante à sa soirée chez des amis. Elle essaye plusieurs robes de son armoire, mais rien ne lui convient. Elle remet la tenue qu'elle portait aujourd'hui.

Nom:  
Prénom:  
Age:  
Classe:

## TACHE B.1

Lisez attentivement le texte suivant.

En sortant du restaurant, les parents de Théo veulent prendre un raccourci pour rentrer chez eux. Après avoir marché pendant plus de 15 minutes, les voilà de nouveau devant l'entrée du restaurant !

Nom:  
Prénom:  
Age:  
Classe:

Puis lisez séparément chacun des textes ci-dessous. A chaque fois qu'un texte exprime la même idée que

le texte que du haut, cochez la case correspondante. Vous pouvez cochez plusieurs cases ou aucune.



Yanis a rendez-vous à la piscine avec ses amis. Sa mère le conduit en voiture mais ils sont coincés dans les bouchons. Il descend de la voiture et finit le trajet en courant. Il arrive tout de même à l'heure à son rendez-vous.

Léa est très pressée et décide de rentrer chez elle en bus plutôt qu'en marchant. Elle prend le bus à la station devant la gare. À peine assise dans le bus, Léa s'endort et manque son arrêt. Lorsqu'elle rouvre les yeux, le bus est revenu devant la gare.

Sacha a un problème de physique à résoudre, il ne trouve pas la solution. Depuis plusieurs jours, il travaille à son bureau, mais il n'a rien écrit sur sa feuille. De nombreuses idées lui viennent à l'esprit mais aucune ne fonctionne.

Chloé souhaite se changer pour être plus élégante à sa soirée chez des amis. Elle essaye plusieurs robes de son armoire, mais rien ne lui convient. Enfin, elle porte la tenue qu'elle portait aujourd'hui.

**ANALOGUE NIVEAU 2**

**ANALOGUE NIVEAU 3**

## C. PROCEDURE

Pour chacune des 16 classes, les élèves étaient répartis en deux demi-groupes afin de permettre un éparpillement « un par table » des élèves dans la classe et d'éviter autant que possible une influence entre voisins. La passation des expériences se déroulait successivement pour chaque demi-groupe dans la salle de classe pendant une durée de 45 minutes (l'autre moitié de la classe étant prise en charge par un autre professeur). Cette durée permettait d'établir le contact avec les élèves, d'expliquer le déroulement de la séance, la réalisation des différentes tâches ainsi qu'une période de retour d'expérience de 10 minutes avec les élèves.

Au cours du préambule oral, les élèves étaient invités à prendre tout le temps dont ils avaient besoin (sous réserve de ne pas dépasser 15 minutes par tâche) puisqu'il ne s'agissait pas d'une épreuve de vitesse. De plus, il leur était précisé qu'il ne s'agissait pas d'un test, qu'il n'y avait pas de bonnes ni de mauvaises réponses et qu'ils devaient fournir les réponses qui leur semblaient les plus pertinentes.

Les consignes étaient données ensuite à l'oral au début de chaque tâche et distribuées sous forme écrite (voir p 31 et 34). Il était précisé aux élèves que ni l'expérimentateur ni le professeur ne pourraient répondre à aucune question une fois la consigne donnée. Nous noterons néanmoins l'intervention des professeurs auprès de trois élèves en ULIS présentant de grosses difficultés de lecture. Notre objectif étant de mesurer la capacité à percevoir des analogies entre des textes qu'ils soient lus ou racontés, nous avons décidé d'inclure ces résultats notamment étant donné le petit nombre d'effectifs concernés.

La séance consistait en la réalisation de deux tâches distinctes, exécutées plusieurs fois chacune avec des jeux de stimuli distincts (deux fois pour la tâche 1, trois fois pour la tâche 2). Un contrebalancement a été introduit de manière à ce que chaque stimulus apparaisse aussi fréquemment dans chacun des ordres possibles et qu'entre voisins les élèves n'exécutent pas les mêmes versions simultanément. Les deux tâches étaient toujours exécutées dans le même ordre en commençant par la tâche de catégorisation.

A la fin de chaque séance, les dix dernières minutes étaient consacrées à un retour d'expérience des élèves quant aux difficultés rencontrées, à leurs opinions sur les tâches qu'ils venaient de réaliser. De plus trois élèves par demi-groupe étaient interrogés individuellement par l'examinateur pour donner des détails sur leurs réponses. Certains élèves choisis pour ces entretiens étaient ceux dont des réponses pour la tâche 2 avaient été inattendues (non perception de l'analogie de niveau 1 ou non perception de celle de niveau 2 alors que celle de niveau 3 avait été perçue). L'objectif de ce retour d'expérience était de relever toute difficulté ou problème perçu par les élèves aussi bien de manière groupée qu'individuel ; les élèves ayant parfois peur de signaler des difficultés devant l'ensemble de la classe. Aucun problème particulier n'a été relevé lors de ces sessions et nous a permis de recueillir des explications (oubli, inattention, impression de piège) sur certains cas inattendus de non perception de l'analogie de niveau 1.

## B. ANALYSE QUANTITATIVE

Dans un premier temps, les réponses données par chaque élève ont été analysées séparément afin de mettre en évidence la logique générale de classement adoptée par chaque individu ; dans un deuxième temps, ces logiques individuelles de réponse ont été rassemblées et comparées par populations d'enfants identifiées a priori. Plusieurs mesures quantitatives ont été introduites afin de mettre en évidence les résultats obtenus lors de chaque version des deux tâches et tenter de répondre à notre hypothèse.

### **Tâche 1 :**

Nous avons défini tout d'abord pour chaque élève des variables de Bernoulli **SS** et **SP**, qui prennent la valeur 1 lorsque la classification de ce dernier matche parfaitement la catégorisation hypothétique correspondante, et 0 lorsque ce n'est pas le cas. Nous avons calculé ensuite **les distances de correction** de sa réponse par rapport à chacune des deux catégorisations hypothétiques. Nous les nommerons respectivement **DC\_SS** et **DC\_SP**. La distance de correction par rapport à une catégorisation hypothétique se définit comme le nombre minimum de mouvements nécessaires pour transformer la catégorisation du sujet en cette

dernière. Supposons par exemple que la catégorisation hypothétique soit {ABC, DEF, GHI} et que la réponse de l'élève soit {AB, CEF, DGHI}, il faut deux mouvements pour passer de l'une à l'autre. Avec des jeux de 9 cartes et une catégorisation hypothétique de trois fois trois cartes, la plus grande distance de correction qui puisse être attribuée à un élève est donc de 6. Plus cette distance de correction est petite, plus cela signifie que la catégorisation de l'élève est proche de la catégorisation hypothétique considérée.

Nous nous sommes intéressés aussi aux nombres de paires communes entre la classification de l'élève et les catégorisations hypothétiques afin de déterminer le nombre de paires (SF), de paires (SS) ainsi que celles inattendues pour chaque sujet. **Les taux d'appariement (PA\_SF, PA\_SS, PA\_SI)** ont été obtenus en divisant ces nombres par le nombre total de paires possibles dans la réponse du sujet. Si nous reprenons notre exemple précédent dans lequel les catégorisations hypothétiques (SS) et (SP) sont respectivement {ABC, DEF, GHI} et {ADG, BEH, CFI}, les taux d'appariement associés à la réponse {AB, CEF, DGHI} sont donc  $PA_{SS} = 5/10$  ;  $PA_{SP} = 2/10$   $PA_{SI} = 3/10$ . Plus la réponse de l'élève est proche d'une des catégorisations hypothétiques, plus le pourcentage d'appariement correspondant sera élevé.

Nous avons analysé enfin les regroupements de cartes récurrents au sein des réponses des élèves, ainsi que leurs contenus. Des scores de similarité ont été calculés pour chaque paire de cartes possible, basés sur la fréquence avec laquelle cette paire a été classée dans la même catégorie au sein d'une de nos sous sous-populations élèves. Ces scores ont été réunis dans une matrice de similarité, de manière à ce que pour toute paire de cartes  $[i,j]$ , la cellule de la matrice correspondant à la  $i^{\text{ème}}$  ligne de la  $j^{\text{ème}}$  colonne contienne le score de similarité entre  $i$  et  $j$ . Ainsi par exemple, si la carte B est classée par 150 participants (sur un total de 407) avec la carte F, le score de similarité situé dans la cellule (6,2) de la matrice sera égal à  $150/407 \times 100 = 37\%$ . Plusieurs matrices de similarités ont ainsi été créées, représentant chacune un des sous-groupes d'élèves. Sur la base de ces matrices, des dendrogrammes ont été construits pour illustrer visuellement le lien de proximité entre les différentes cartes pour chaque sous-groupe et ainsi faciliter l'interprétation des résultats.

Les titres donnés par les élèves aux différentes catégories n'ont pas été codés ni analysés dans ce travail. Cette consigne avait pour premier objectif de maximiser l'engagement de l'enfant, d'éviter qu'il ne réponde de manière aléatoire et de vérifier la cohérence de chaque réponse dans la phase de codage. Ces mesures ainsi définies, nos hypothèses ont été déclinées de manière opérationnelle pour la tâche 1 de la façon suivante :

**Hypothèse 1.A :** Nous nous attendons à ce que la proportion de classifications (SP) basées sur des traits de structure versus celles basées sur des éléments de surface (SP) dans chaque sous-groupe d'élèves varie crescendo avec le niveau scolaire du groupe. Nous anticipons des mesures DC\_SS et PA\_SP plus élevées en moyenne pour les élèves de niveau A que pour ceux de niveau B (et avec la même logique, plus élevées pour ceux de niveau B que pour ceux de niveau C). Inversement, les mesures DC\_SP et PA\_SS observées devraient être plus basses en moyenne pour les élèves de niveau A que pour ceux de niveau B (et avec la même logique, plus basses pour ceux de niveau B que pour ceux de niveau C).

**Hypothèse 1.B :** Nous nous attendons à observer un plus grand nombre de classifications basées sur des traits de structure dans le groupe HP que dans le groupe Standard. Ce qui devrait se traduire par une proportion de succès pour la variable SP et des mesures DC\_SS et PA\_SP plus élevées en moyenne pour les élèves HP que pour les Standard, ainsi qu'une proportion d'échecs de la variable SS et des mesures DC\_SP et PA\_SS plus faibles.

## **Tâche 2 :**

Afin de quantifier les différences en termes de niveaux d'abstraction dont font preuve les élèves, nous avons calculé les taux d'analogie reflétant les proportions de textes cibles d'un certain niveau d'abstraction ayant été associés à un texte source donné. Pour se faire nous avons introduit pour chaque version de la tâche 2 :

- des variables de Bernouilli **Niv1**, **Niv2** et **Niv3** qui prennent la valeur 1 lorsque l'analogie du niveau correspondant est perçue par l'élève, 0 dans le cas contraire ;
- des variables de Bernouilli **Intrus**, **Erreur** qui prennent respectivement la valeur 1 lorsque l'intrus est sélectionné par l'élève ou lorsque l'analogie de niveau 1 n'est pas perçue par l'élève, 0 dans les cas contraires ;
- la variable polyatomique ordinaire **Score**, qui synthétise la performance de l'élève dans cet exercice. Les valeurs, que prennent cette variable Score ont été définies à l'aide du tableau 4.3 ci-dessous, reflétant le classement que nous avons établi en amont des différents scénarios de réponse. Ce classement équivaut à l'attribution d'un point pour l'analogie de niveau 1, de deux points pour celle de niveau 2, et de quatre points pour celles de niveau 3. La sélection de l'intrus entraîne quant à elle la perte d'1/2 point.

Une fois ces métriques définies, nous avons formulé les hypothèses opérationnelles suivantes pour la tâche 2:

**Hypothèse 2.A:** Nous anticipons une corrélation entre le niveau scolaire des élèves et leur capacité à percevoir des structures communes à un niveau élevé d'abstraction et par conséquent nous nous attendons à observer des taux d'analogie de niveau 1, niveau 2 et niveau 3 (respectivement des taux d'erreur et d'intrus) plus élevés (respectivement plus faibles) dans le groupe de niveau A que dans celui de niveau B avec un effet plus significatif pour le niveau d'abstraction le plus élevé. Nous prédisons un phénomène similaire entre les groupes B et C.

**Hypothèse 2.B :** Nous nous attendons à ce que les enfants HP perçoivent les relations de similarités à un niveau d'abstraction plus élevé que les autres enfants, et par conséquent que les taux d'analogie de niveau 1, niveau 2 et niveau 3 (resp. les taux d'erreur et d'intrus) soient plus élevés (resp. plus faibles) dans le groupe HP que dans le groupe standard, avec un effet plus significatif pour le niveau d'abstraction le plus élevé.

Tableau 4.3: Définition de la variable score en fonction  
des similarités perçues

<b>Textes Sélectionnés</b>	<b>Score</b>	<b>Rang</b>
Niv1, Niv2, Niv3	7	1
Niv1, Niv2, Niv3, Intrus	6.5	2
Niv2, Niv3	6	3
Niv2, Niv3, Intrus	5.5	4
Niv1, Niv3	5	5
Niv1, Niv3, Intrus	4.5	6
Niv3	4	7
Niv1, Niv2	3	8
Niv1, Niv2, Intrus	2.5	9
Niv 2	2	10
Niv2, Intrus	1.5	11
Niv 1	1	12
Niv1, Intrus	0.5	13
Aucun	0	14
Intrus seul	-0.5	15

## V. Résultats

---

Nous avons mené une série d'analyses préliminaires pour chacune des tâches, ayant pour but d'évaluer l'impact des différents variables susceptibles d'influencer nos résultats et donc de s'assurer de la pertinence de notre dispositif expérimental pour tester nos hypothèses. Pour chacune des variables dépendantes associées à nos deux tâches et en fonction de sa nature, des analyses univariées (test du khi-carré, exact de Fisher ou de Student) et multivariées (régression logistique, ordinaire ou ANOVA) ont été réalisées tenant compte des facteurs que nous avons identifiés comme pouvant expliquer les différences observées entre nos sous-groupes d'élèves et pour lesquels nous disposons des données.

Nous nous sommes ainsi intéressés aux facteurs suivants : certains intrinsèques à l'enfant (Age, Sexe, Lycée et Classe) d'autres propres au dispositif expérimental (Version des stimuli, Ordre de passation) en addition de nos variables d'intérêt représentés par le Niveau Scolaire et le Diagnostic de précocité - ainsi qu'à toute sorte d'interactions qui pouvaient exister entre eux. Une fois clarifiée l'effet des différents facteurs et les potentielles sources de confusion, la suite des analyses a porté sur la comparaison des réponses des élèves d'une part suivant leur niveau de performance scolaire et d'autre part suivant le fait qu'ils aient été diagnostiqués comme précoces.

Les réponses des élèves ont été codées en Matlab afin d'extraire pour chaque sujet les variables dépendantes considérées. Les analyses ont toutes été menées à l'aide du logiciel Stata IC15 et pour chaque analyse les conditions d'application (tels l'indépendance des observations, la normalité des résidus, l'homogénéité de la variance etc...) ont été contrôlées. Pour la quasi-totalité de nos variables dépendantes nous avons observés des violations de l'hypothèse de normalité (rejet de l'hypothèse nulle au test de Shapiro Wilk) et des violations de l'hypothèse d'homogénéité de la variance (rejet de l'hypothèse nulle au test de Levene) principalement pour les sous-groupes HP et Standard, sous-groupes

de taille largement distinctes. Les modèles ANOVA étant souvent robustes aux déviations de l'hypothèse de normalité des Welch ANOVA ont été utilisés pour palier au problème de non-homogénéité de la variance. En parallèle, des tests non paramétriques ont été réalisés dans l'analyse principale de manière systématique pour s'assurer de la validité de nos conclusions en dépit des violations des hypothèses. Il s'agit des tests de Mann-Whitney U et de Kruskall-Wallis H en fonction du nombre d'échantillons que nous comparons. L'ensemble des résultats obtenus suivant les différentes méthodes seront présentées. Les conclusions de l'ensemble de ces tests étant cohérents avec des p values proches de ceux obtenus avec nos modèles ANOVA à un facteur, c'est pourquoi nous avons fait le choix de conserver nos modèles paramétriques multivariées comme modèle pour les analyses préliminaires des cinq variables dépendantes non binaires malgré les violations de normalité.

## A. TACHE 1 DE CATÉGORISATION

Le plan expérimental pour cette première tâche est un plan expérimental non équilibré avec plusieurs variables dépendantes :

- les variables de Bernoulli SS, SP reflétant le match parfait avec l'une des deux catégorisations hypothétiques;
- les distances de correction DC\_SS, DC\_SP par rapport à nos catégorisations hypothétiques;
- les taux d'appariement PA\_SS, PA\_SP, PA\_SI (pourcentages des paires communes entre la réponse de l'élève et chacune des 2 catégorisations hypothétiques ainsi que le pourcentage de paires inattendues, i.e. présentes dans aucune des deux);

et les variables indépendantes suivantes: Niveau scolaire, Diagnostic de précocité, Lycée, Classe, Age, Sexe, Version des stimuli, Ordre de passation.

## 1. ANALYSES PRELIMINAIRES

### Effets liés au dispositif expérimental : modalité et ordre de présentation des stimuli

Nous présentons dans ce travail les réponses des élèves pour une seule réalisation de la tache 1. Étant donné que les deux versions de la tache ont toutes été exécutées par les élèves mais avec un ordre aléatoire, deux axes d'études auraient été possibles : soit nous considérions toutes les réalisations de la version A des stimuli, quel que soit l'ordre de présentation aux élèves, soit nous nous intéressions à l'ensemble des premières réponses fournies par les élèves, quelle que soit la modalité des stimuli utilisés. Il est en effet légitime de se demander si d'une part des phénomènes d'apprentissage ou inversement d'inattention se sont manifestés lors des résolutions successives de la même tache, et si d'autre part bien que les deux versions de nos stimuli aient été élaborées de la même façon, des différences liées aux contextes ou aux textes eux-mêmes sont apparues.

Des analyses univariées et multivariées à mesures répétées ont été menées à cet effet sur l'ensemble formé de nos 2 jeux de données pour la tache 1 (test du khi-carré et régressions logistiques pour les variables SS et SP ; ANOVA pour les cinq variables DC\_SS, DC\_SP, PA\_SS, PA\_SP, et PA\_SI) pour étudier les effets potentiels des différentes modalités et ordres de passation sur les proportions/ moyennes des variables dépendantes. Les p values (ajustées et non ajustées) obtenus pour chaque facteur et chaque variable dépendante étaient à chaque fois proches et non significatives ( $p > 0.20$ ).

Nous avons mené ces analyses préliminaires bien qu'ayant pris soin de contrebalancer l'ordre des différents stimuli de manière aléatoire lors de la passation de l'expérience dans chaque classe. Car ne connaissant pas les caractéristiques intrinsèques à chaque élève lors de la distribution des énoncés nous n'avons pu contrebalancer l'ordre de présentation au sein des différents sous-groupes que nous souhaitions étudier (voir le tableau 5.A.2).

		Tache 1	
Ordre des stimuli		A-B	B-A
Echantillon 1	HP	29	13
	Non-HP	70	89
	Total	99	102
Echantillon 2	A	31	36
	B	34	42
	C	37	25
	Total	102	103

Tableau 5.A.1: Repartition de l'ordre de passation des stimuli pour la tache 1 suivant les sous-groupes d'élèves

De plus, afin de s'assurer de l'homogénéité des mesures répétées pour chaque variable dépendante sur l'ensemble des sujets et ce de manière non paramétrique, nous avons réalisés des tests de McNemar (pour les variables binaires SS et SP) et des tests de Friedman (pour les variables DC\_SS, DC\_SP, PA\_SS, PA\_SP, et PA\_SI). Les p values obtenues n'étaient pas significatives et n'ont pas permis de rejeter l'hypothèse nulle de non effet entre les deux passations.

Les analyses statistiques ayant montré que ni l'ordre de passation ni la version des stimuli n'influaient nos résultats de manière significative, nous avons choisi de considérer comme jeu d'observations indépendantes, dans un premier temps uniquement l'ensemble des réalisations de la version A des stimuli, quel que fut son ordre de présentation aux élèves (puis dans un deuxième temps uniquement ceux correspondant à la version B). La raison étant de faciliter le codage ainsi que de maximiser la cohérence des matrices de proximité et des dendrogrammes en ne basant leur construction que sur le même lot d'histoires. Les résultats étant équivalents pour les deux versions nous ne détaillerons ici que ceux de la version A.

### **Effets liés aux caractéristiques des élèves : Age, Sexe, Lycée et Classe**

La deuxième partie de nos analyses préliminaires s'est portée sur l'étude des effets des variables Age et Sexe (seules données intrinsèques à l'élève dont nous disposions en addition de la langue maternelle). En effet, comme l'illustre le tableau 5.A.1, ces variables ne sont pas contrebalancées entre les différents sous-groupes de nos deux échantillons. Des séries d'analyses univariées et multivariées ont été conduites pour chaque facteur sur chacune des variables dépendantes afin de voir s'il y avait des effets significatifs sur leurs moyennes

ou proportions. Seuls les effets liés au facteur âge se sont révélés significatifs dans les analyses univariées, mais il s'agissait d'un effet de confusion lié aux facteurs Niveau scolaire et Diagnostic de précocité, et dans nos modèles multivariés toutes les p values ajustées associées au facteur âge étaient  $p > 0,10$ . En effet, les enfants les plus âgés dans notre échantillon total sont ceux dont les résultats scolaires sont les moins bons et réciproquement nous trouverons une stricte majorité d'enfants diagnostiqués HP ou de niveau A chez ceux de 10 ans (voir tableau 4.A.1).

Nous avons aussi inclus dans nos modèles multivariés la variable Classe pour les données provenant de l'échantillon 1 et les variables Classe et Lycée pour les données provenant de l'échantillon 2 (l'échantillon 1 n'étant constitué que des élèves du lycée 1) ainsi que les potentielles interactions entre les deux. Aucune différence significative n'a été révélée pour aucun de ces facteurs.

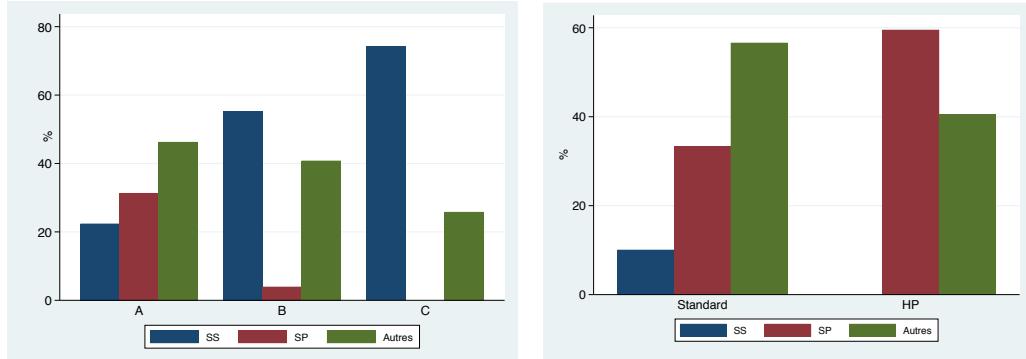
Si dans les analyses préliminaires nous avons voulu caractériser les relations entre les variables dépendantes et l'ensemble des différentes variables explicatives, nous ne nous concentrons plus que sur les effets des facteurs Niveau scolaire et Diagnostic de précocité dans la suite de ce travail mais nos modèles multivariés (régression logistique ou ANOVA suivant la nature de la variable dépendante) mentionnés dans les analyses principales intègrent l'ensemble des différents facteurs afin de minimiser les effets de confusion.

## 2. ANALYSES PRINCIPALES

Afin d'évaluer la validité de nos hypothèses 1.A et 1.B nous avons menées deux études comparatives distinctes : l'une sur les différents groupes de niveau de l'échantillon 2 d'une part, l'autre sur les groupes HP et Standard de l'échantillon 1 d'autre part.

Les graphes et tableaux 5.A.2 illustrent les proportions de réponses des élèves correspondant parfaitement aux deux catégorisations hypothétiques ainsi que celles ne correspondant pas, pour les différents sous-groupes des deux échantillons.

**Fig. 5.A.2:** Comparaison des proportions de catégorisations hypothétiques de surface (SS) et de structure (SP) réalisées par les élèves, suivant les niveaux scolaires A, B ou C de l'échantillon 2 (à gauche), suivant les groupes Standard vs. HP de l'échantillon 1 (à gauche).



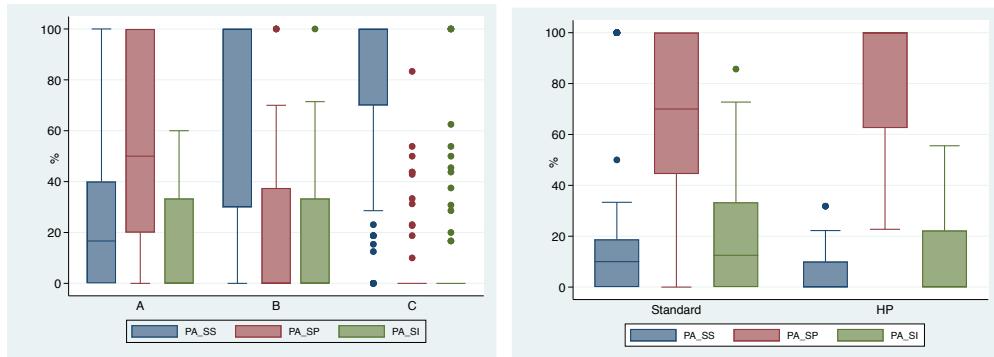
Niveau Scolaire	Eleves N	Proportions de match exact %: Moy. et IC				Type	Eleves N	Proportions de match exact %: Moy. et IC			
		(SS)	(SP)	(SS)	(SP)			(SS)	(SP)	(SS)	(SP)
A	67	22	13	34	31	21	44	10	6	16	33
B	76	55	43	67	4	1	11	0	0	8	43
C	62	74	62	84	0	0	6	159	26	41	74
Total	205	50	43	57	12	8	17	42	39	32	46
							Total	201	8	5	13

Conformément à nos hypothèses, nous observons un plus grand nombre de classifications (SP) chez les élèves avec le meilleur niveau scolaire ainsi que dans le groupe HP versus celui standard. Nous remarquons l'absence de catégorisation (SP) dans le groupe de niveau C au sein de l'échantillon 2 ainsi que l'absence de catégorisation (SS) dans le groupe HP au sein de l'échantillon 1. Les tests statistiques que nous avons menés confirment la significativité des différences observées. Les tableaux 5.A.5 et 5.A.6 détaillent les résultats obtenus suite aux régressions logistiques et aux tests non-paramétriques du khi-carré/exact de Fischer sur les variables binaires SS et SP. Les odd ratios fournissent des informations sur la force et le sens de l'association entre nos variables dépendantes et les facteurs Niveau scolaire et Diagnostic de précocité.

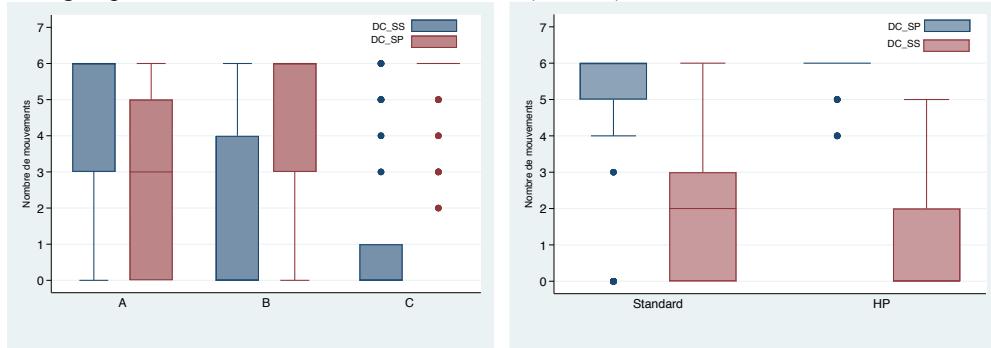
Ces résultats nous révèlent aussi le grand nombre de classifications ne correspondant à aucune des deux catégorisations hypothétiques dans chaque sous-groupe des deux échantillons. Cette proportion est particulièrement élevée au sein de l'échantillon 1. Ces réponses correspondent à des critères de classification basés à la fois sur des éléments de surface et de structure mais aussi parfois sur des similarités perçues propres à l'encodage de l'élève.

Afin d'évaluer le niveau d'abstraction des critères utilisés par les élèves des différents groupes, nous nous sommes intéressés aux pourcentages de paires communes entre la réponse d'un élève et les deux catégorisations hypothétiques ainsi qu'au pourcentage de paires inattendues. Nous avons aussi mesuré les distances de correction, i.e. le nombre minimal de mouvements nécessaires pour passer de la catégorisation de l'élève à l'une des catégorisations hypothétiques. Les graphes 5.A.2 et les tableaux 5.A.3 détaillent les mesures moyennes obtenues (ainsi que les erreurs standards) sur les sous-groupes des deux échantillons. Pour normaliser les différences de tailles respectives des sous-groupes considérés, les variances sont présentées en termes d'erreur standard à la moyenne.

**Fig. et Tab 5.A.3:** Comparaison des pourcentages moyens de paires communes avec les catégorisations (SS) et (SP) ainsi que des paires inattendues des groupes de niveau A, B ou C de l'échantillon 2 (à gauche), et des groupes Standard vs. HP de l'échantillon 1 (à droite).



**Fig. et Tab 5.A.4:** Comparaison des distances moyennes de correction par rapport aux catégorisations (SS) et (SP) des groupes de niveau A, B ou C de l'échantillon 2 (à gauche), des groupes Standard vs. HP de l'échantillon 1 (à droite).



Niveau Scolaire	Eleves	Distance de correction DC :				% PA de paires communes:					
		Moyenne (Err. St.)				Moyenne (Err. St.)					
		N	(SS)	(SP)		(SS)	(SP)	(SI)			
A	67	4.3	(0.30)	2.6	(0.28)	30	(5)	55	(5)	14	(2)
B	76	2.0	(0.28)	4.7	(0.21)	66	(5)	19	(3)	15	(3)
C	62	1.2	(0.15)	5.5	(0.13)	80	(5)	7	(2)	13	(4)
<b>Total</b>	<b>205</b>	<b>2.5</b>	<b>(0.19)</b>	<b>4.3</b>	<b>(0.15)</b>	<b>58</b>	<b>(3)</b>	<b>27</b>	<b>(2)</b>	<b>14</b>	<b>(2)</b>

Type	Eleves	DC Distance de correction :				% PA de paires communes:					
		Moyenne (Err. St.)				Moyenne (Err. St.)					
		N	(SS)	(SP)		(SS)	(SP)	(SI)			
<b>Standard</b>	159	4.9	(0.15)	2.0	(0.15)	19	(2)	65	(3)	16	(1)
<b>HP</b>	42	5.7	(0.09)	1.0	(0.22)	5	(1)	83	(4)	12	(3)
<b>Total</b>	<b>201</b>	<b>5.1</b>	<b>(0.11)</b>	<b>1.8</b>	<b>(0.13)</b>	<b>15</b>	<b>(2)</b>	<b>69</b>	<b>(2)</b>	<b>15</b>	<b>(1)</b>

Plus la classification d'un élève se base sur des éléments de structure et est proche de la catégorisation hypothétique SP, plus le pourcentage PA\_SP est élevé et se rapproche de 1 et plus la distance de correction DC\_SP est petite et se rapproche de 0. Le pourcentage PA\_SS et la distance de correction DC\_SS quant à eux évoluent de façon contraire : PA\_SP se rapproche de 0 et DC\_SS de 6, qui comme nous l'avons vu précédemment correspondant au nombre de mouvements maximum pour atteindre une catégorie donnée dans tout jeu de 9 cartes.

Afin de confirmer les différences significatives observées, nous avons réalisé des ANOVA multifactoriels et Welch ANOVA (à un facteur) suivis des post hoc test de Games-Howell ainsi que les tests non-paramétriques de Mann-Whitney U et de Kruskall-Wallis H suivis des post hoc test de Dunn sur nos variables dépendantes DC et PA. Les tableaux 5.A.5 et 5.A.6 détaillent les résultats obtenus. Les p values significatives (en tenant compte de l'ajustement Bonferroni pour les comparaisons multiples) apparaissent en gras. Les différents tests s'accordent pour confirmer la significativité des effets observés dans nos deux échantillons, à l'exception des pvalues non significatives associées au pourcentage de paires inattendues dans la réponse des élèves.

**Tableau 5.A.5:** Tache 1.A - Statistiques, effets et p values associés aux réponses des élèves de l'échantillon 2 selon les groupes de niveaux scolaires A, B ou C.

Statistiques	SS	SP	DC_SS	DC_SP	PA_SS	PA_SP	PA_SI
<b>Modele parametrique</b>							
<b>Modele</b>	Reg.Logistique $z= 5.37$ $p=0.000$	Reg.Logistique $z = -4.23$ $p=0.000$	ANOVA <b>F(2,194)=28.27</b> <b>p=0.000</b>	ANOVA <b>F(2,194)=43.6,</b> <b>p=0.000</b>	ANOVA <b>F(2,194)=26.44,</b> <b>p=0.000</b>	ANOVA <b>F(2,194)=46.2,</b> <b>p=0.000</b>	ANOVA $F(2,194)=0.02,$ $p=0.976$
<b>Stat/ pvalue</b>							
<b>Modele</b>			Welch ANOVA <b>W(2,133.44)=29.8</b>	Welch ANOVA <b>W(2,125.83)=44.9</b>	Welch ANOVA <b>W(2,133.62)=28.97,</b> <b>3, p=0.000</b>	Welch ANOVA <b>W(2,126.11)=42.69,</b> <b>6, p=0.000</b>	Welch ANOVA $W(2,127.00)=0.10,$ $p=0.902$
<b>Stat/ pvalue</b>							
<b>Effet (SEM) pvalue</b>							
B vs. A	OR: 4.47 (1.72), $p=0.000$	OR: 0.08 (0.05), $p=0.000$	-2.24 (0.641) $p=0.000$	2.07 (0.35), $p=0.000$	35.6 (6.7), $p=0.000$	-36.1 (5.68), $p=0.000$	Non Sig
C vs. A	OR: 9.62 (4.07), $p=0.000$	OR: 1	-3.06 (0.41) $p=0.000$	2.91 (0.31), $p=0.000$	49.2 (6.6), $p=0.000$	-47.81 (5.2), $p=0.000$	Non Sig
C vs. B	OR: 2.15 (0.81), $p=0.042$	OR: 1	Non sig	0.84 (0.25) $p=0.003$	Non sig	-11.74 (3.9), $p=0.009$	Non sig
<b>Modele non-parametrique</b>							
<b>Modele</b>	chi2	chi2	Kruskall Wallis H	Kruskall Wallis H	Kruskall Wallis H	Kruskall Wallis H	Kruskall Wallis H
	<b>chi2(2) = 35.79,</b> $p=0.000$	<b>chi2(1) = 37.6,</b> $Fischer p=0.000$	<b>chi2(2) = 48.66,</b> $p=0.000$	<b>chi2(2) = 59.44,</b> $p=0.000$	<b>chi2(2)= 40.08,</b> $p=0.001$	<b>chi2(2)= 61.48,</b> $p=0.001$	<b>chi2(2)= 3.521,</b> $p=0.172$
<b>Stat/ pvalue</b>							
<b>Effet (SEM) pvalue</b>							
B vs. A	-32.9 (7.7) $p=0.000$	27.4 (6.1) $p=0.000$	$p=0.000$	$p=0.000$	$p=0.000$	$p=0.000$	Non sig
C vs. A	-51.9 (7.5) $p=0.000$	31.3 (5.7) $p=0.000$	$p=0.000$	$p=0.000$	$p=0.000$	$p=0.000$	Non sig
C vs. B	-18.9 (8.0) $p=0.021$	Non Sig	Non Sig	$p=0.022$	Non sig	$p=0.031$	Non sig

**Tableau 5.A.6:** Tache 1.A - Statistiques, effets et p values associés aux réponses des élèves de l'échantillon 1 selon les groupes Standard et HP.

Statistiques	SS	SP	DC_SS	DC_SP	PA_SS	PA_SP	PA_SI
<b>Modele parametrique</b>							
<b>Modele</b>	Reg.Logistique predict perf. $z = 2.03$ $p=0.042$	Reg.Logistique $z = -4.23$ $p=0.000$	ANOVA <b>F(1,193)=6.8,</b> <b>p=0.010</b>	ANOVA <b>F(1,193)=6.1,</b> <b>p=0.015</b>	ANOVA <b>F(1,193)=7.83,</b> <b>p=0.006</b>	ANOVA <b>F(1,193)=7.24,</b> <b>p=0.008</b>	ANOVA $F(1,193)=0.45,$ $p=0.5$
<b>Stat/ pvalue</b>							
<b>Modele</b>			Welch ANOVA <b>W(1, 185)=19.51,</b> $0.79 (0.3), p=0.008$	Welch ANOVA <b>W(1,84.18)=12.7,</b> $-1.0 (0.3) p=0.003$	Welch ANOVA <b>W(1,195.6)=24.75,</b> $-13.7 (4.7), p=0.004$	Welch ANOVA <b>W(1,88.56)=15.20,</b> $17.7 (5.5), p=0.002$	Welch ANOVA $W(1,65.72)=1.72,$ $p=0.170$
<b>Stat/ pvalue</b>							
<b>Effet (SEM) pvalue</b>							
B vs. A	10.1 (2.4)	OR: 2.19 (0.84)	Mann-Whitney U <b>chi2(1) = 4.59,</b> $Fischer p=0.027$	Mann-Whitney U <b>chi2(1) = 9.60,</b> $p=0.002$	Mann-Whitney U <b>chi2(1)=11.30,</b> $p=0.001$	Mann-Whitney U <b>chi2(1)=10.97,</b> $p=0.001$	Mann-Whitney U $chi2(1)= 1.92$
C vs. A	10.1 (2.4)	-26.2 (8.5)	$p=0.005$	$p=0.0017$			
C vs. B							

Pour conclure notre étude de cette première tâche, nous avons voulu utiliser de plus des analyses ne faisant pas référence à des catégorisations hypothétiques. Ainsi les classifications réalisées spontanément par les élèves ont été retranscrites à l'aide de matrices de cooccurrence indiquant quels textes ont été regroupés le plus fréquemment ensemble au sein d'une même groupe d'élèves. Les matrices ci-dessous représentent les données de notre échantillon 2 réparties entre les différents groupes de niveaux scolaires et celles de l'échantillon 1 réparties entre les groupes HP et Standard. Rappelons ici que la catégorisation hypothétique (SS) se traduit par les groupes {ACF} {BGH} {DEI} et la catégorisation hypothétique (SP) par {ABD} {CEG} {FHI}.

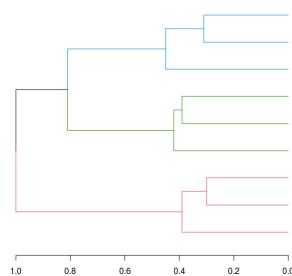
Groupe de Niveau A - Echantillon 2										< 20% 21 - 50 % > 50%		
Cartes	A	B	C	D	E	F	G	H	I			
A	100%	61%	24%	61%	1%	48%	7%	25%	25%			
B	61%	100%	1%	58%	0%	21%	27%	46%	19%			
C	24%	1%	100%	0%	70%	25%	61%	1%	0%			
D	61%	58%	0%	100%	25%	31%	1%	22%	49%			
E	1%	0%	70%	25%	100%	3%	61%	0%	24%			
F	48%	21%	25%	31%	3%	100%	6%	55%	63%			
G	7%	27%	61%	1%	61%	6%	100%	24%	1%			
H	25%	46%	1%	22%	0%	55%	24%	100%	69%			
I	25%	19%	0%	54%	24%	63%	1%	69%	100%			

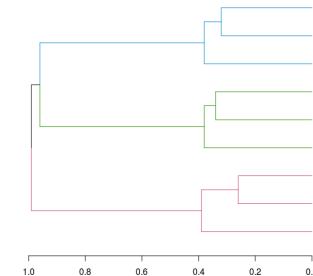
Groupe de Niveau B - Echantillon 2										< 20% 21 - 50 % > 50%		
Cartes	A	B	C	D	E	F	G	H	I			
A	100%	14%	61%	24%	4%	74%	7%	9%	13%			
B	14%	100%	7%	14%	9%	1%	66%	66%	5%			
C	61%	7%	100%	8%	21%	66%	20%	1%	3%			
D	24%	14%	8%	100%	68%	11%	4%	13%	66%			
E	4%	9%	21%	68%	100%	5%	22%	5%	62%			
F	74%	1%	66%	11%	5%	100%	7%	16%	24%			
G	7%	66%	20%	4%	22%	7%	100%	62%	7%			
H	9%	66%	1%	13%	5%	16%	62%	100%	24%			
I	13%	5%	3%	66%	62%	24%	7%	24%	100%			

Groupe de Niveau C - Echantillon 2										< 20% 21 - 50 % > 50%		
Cartes	A	B	C	D	E	F	G	H	I			
A	100%	5%	82%	11%	5%	84%	6%	5%	6%			
B	5%	100%	3%	10%	2%	3%	81%	87%	6%			
C	82%	3%	100%	5%	13%	82%	6%	6%	5%			
D	11%	10%	5%	100%	81%	8%	2%	8%	84%			
E	5%	2%	13%	81%	100%	10%	8%	5%	81%			
F	84%	3%	82%	8%	10%	100%	6%	5%	15%			
G	6%	81%	6%	2%	8%	6%	100%	79%	6%			
H	5%	87%	6%	8%	5%	5%	79%	100%	10%			
I	6%	6%	5%	84%	81%	15%	6%	10%	100%			

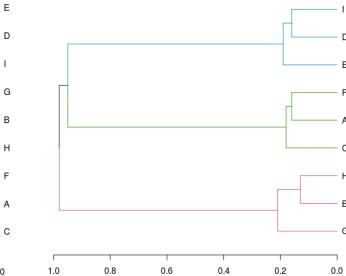
Groupe de Niveau A - Echantillon 2



Groupe de Niveau B - Echantillon 2



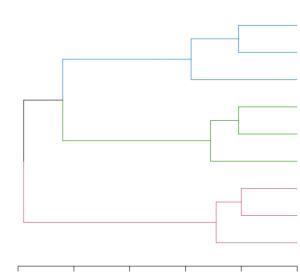
Groupe de Niveau C - Echantillon 2



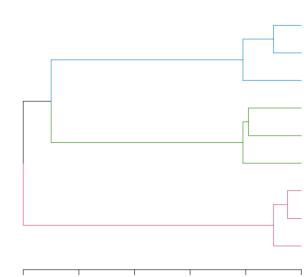
Groupe Standard - Echantillon 1										< 20% 21 - 50 % > 50%		
Cartes	A	B	C	D	E	F	G	H	I			
A	100%	70%	13%	79%	4%	35%	3%	15%	24%			
B	70%	100%	2%	69%	4%	14%	16%	37%	22%			
C	13%	2%	100%	3%	80%	20%	76%	5%	7%			
D	79%	69%	3%	100%	17%	20%	3%	19%	32%			
E	4%	4%	80%	17%	100%	5%	71%	5%	16%			
F	35%	20%	14%	20%	5%	100%	4%	62%	79%			
G	3%	15%	76%	3%	71%	4%	100%	18%	6%			
H	16%	37%	5%	19%	5%	62%	18%	100%	64%			
I	24%	22%	7%	32%	16%	79%	6%	64%	100%			

Groupe HP - Echantillon 2										< 20% 21 - 50 % > 50%		
Cartes	A	B	C	D	E	F	G	H	I			
A	100%	79%	5%	81%	10%	14%	5%	19%	10%			
B	79%	100%	0%	79%	5%	10%	0%	17%	14%			
C	5%	0%	100%	5%	95%	5%	95%	0%	5%			
D	81%	79%	5%	100%	10%	14%	5%	10%	14%			
E	10%	5%	95%	10%	100%	5%	90%	0%	5%			
F	14%	10%	5%	14%	5%	100%	5%	79%	90%			
G	5%	0%	95%	5%	90%	5%	100%	5%	10%			
H	19%	17%	0%	10%	0%	79%	5%	100%	79%			
I	10%	14%	5%	14%	5%	90%	10%	79%	100%			

Groupe Standard - Echantillon 1



Groupe HP - Echantillon 2



Il apparaît clairement que la première dimension ayant guidé la classification des élèves du groupe A est celle basée sur les éléments de structure versus celle basée sur les éléments de surface pour les groupes B et C. Cet effet est aussi visible dans les groupes de l'échantillon 1 mais de façon moins contrastée. Tous les résultats obtenus contribuent donc à confirmer nos hypothèses 1.A et 1.B.

Nous n'avons jusqu'à présent dans cette étude présenté les résultats pour l'échantillon 1 qu'en comparant les données du groupe HP versus celles du groupe Standard. Des effets similaires à ceux obtenus pour l'échantillon 2 ont été observés en comparant les élèves de l'échantillon 1 suivant leur niveau scolaire. Les résultats sont présentés en annexes. Dans nos modèles statistiques (régressions logistiques et ANOVAs) réalisés pour les données de l'échantillon , les facteurs explicatifs les plus forts sont le Diagnostic de précocité et le Niveau Scolaire, et des interactions entre les deux sont de plus présentes. L'existence de ces interactions signifie que l'effet d'une variable est différent en fonction de l'autre variable. Afin d'évaluer l'effet Niveau Scolaire versus Diagnostique HP nous avons effectués de tests de rapport de vraisemblances comparant l'apport du modèle complexe intégrant les deux facteurs versus des modèles tenant compte que de l'un ou l'autre des facteurs. Pour l'ensemble des modèles associés aux six variables dépendantes, les p values obtenues sont à chaque fois significatives et conduisent à rejeter l'hypothèse nulle selon laquelle les variables introduites le modèle complexe n'ont aucun pouvoir explicatif. Il convient de rappeler que nous sommes restreints dans cette analyse par le petit nombre d'élèves à HP de notre échantillon n'ayant pas un niveau scolaire A.

## B. TACHE 2 DE DETECTION D'ANALOGIE

Le plan expérimental pour la deuxième tâche est un plan expérimental non équilibré avec plusieurs variables dépendantes :

- les variables de Bernouilli Niv1, Niv2 et Niv3 codant la réussite de l'identification de l'analogie du niveau correspondant lors de la 1ere passation de la tache ;
- les variables de Bernouilli Intrus et Erreur codant respectivement la sélection du texte intrus et ainsi que la non identification de l'anologue du niveau 1 dans la réponse de l'élève,
- les variables multinomiale TotalNiv1, TotalNiv2 et TotalNiv3 représentant la quantité d'analogies de chaque niveau perçues par l'élève au cours des trois passations ;
- la variable Score, qui permet de noter pour chaque passation la réponse de l'élève par rapport au classement que nous avons établi en amont de tous les scenarii de réponses possibles ;
- la variable TotalScore correspondante à la somme des variables scores associés aux trois passations de la tache 2

et les variables indépendantes suivantes: Niveau scolaire, Diagnostic de précocité, Lycée, Classe, Age, Sexe, Version des stimuli, Ordre de passation.

### 1. ANALYSES PRELIMINAIRES

D'une façon similaire à ce qui a été fait pour la première tâche, nous avons mené des analyses préliminaires pour caractériser les relations entre les variables dépendantes et l'ensemble des variables indépendantes considérées.

Les modèles utilisés et les résultats obtenus sont alignés avec ceux de la tache précédente. Nous mentionnerons néanmoins l'effet significatif lié à la version de l'exercice effectué. En stratifiant les données nous n'avons pas observé de différences significatives associées à l'ordre de passation des différentes versions de stimuli, mais en revanche des effets liés à la version des stimuli lors de chacun des trois passages. Qui plus est, les différences observées ne sont pas identiques suivant les variables réponses considérées. Ainsi par exemple, la version B s'est révélée être la plus réussie pour tous les élèves pour ce qui est de l'identification de l'anologue de niveau 3 (suivie de la C puis de la A), alors que

la version C a montré le plus fort taux d'intrus sélectionné. Comme précédemment bien qu'ayant pris soin de contrebancer l'ordre de distribution des différentes versions dans chaque classe, nous n'avons pu nous en assurer pour chaque groupe d'élèves (voir tableau 5.B.1).

	A-B-C	B-C-A	C-B-A
Echantillon 1	16 51 67	11 55 66	15 53 68
Echantillon 2	34 23 10 67	27 19 30 76	12 22 28 62

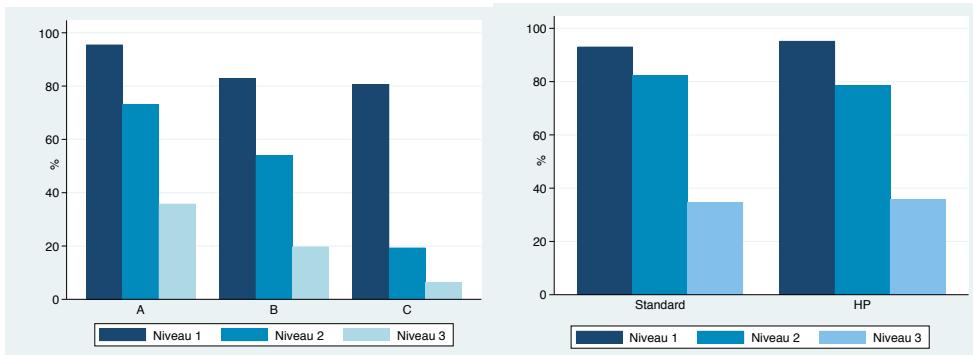
Tableau 5.B.1: Repartition de l'ordre de passation des stimuli pour la tache 2 suivant les sous-groupes d'eleves

En raison de l'effet significatif observé des différentes modalités des stimuli sur les réponses des élèves, les scores de performance global basés sur les réponses de l'élève aux trois passations de la tache ont aussi été analysés. Pour ce faire, les variables dépendantes TotalNiv1, TotalNiv2, TotalNiv3 et TotalScore correspondant à la somme des variables (du même nom) précédemment définies pour chaque passation de la tâche ont été calculés.

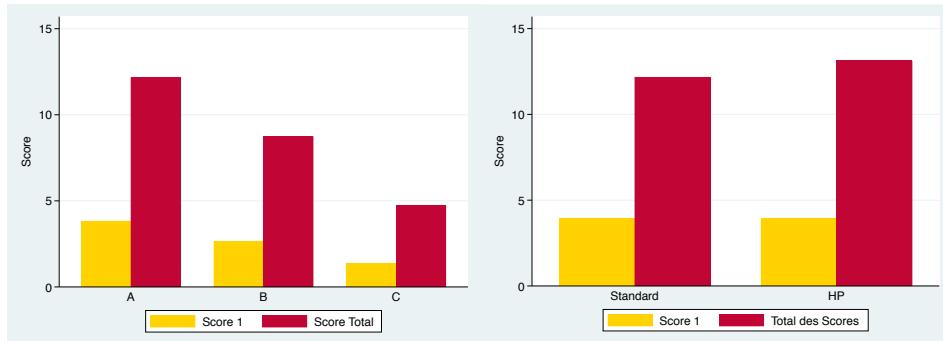
## 2. ANALYSES PRINCIPALES

Suivant la même approche que pour la tâche 1, nous avons mené deux études comparatives distinctes afin d'évaluer la validité de nos hypothèses 2.A et 2.B : l'une sur les différents groupes de niveau de l'échantillon 2 d'une part, l'autre sur les sous-groupes HP et Standard de l'échantillon 1 d'autre part. Les graphes et tableaux 5.B.2, 5.B.3 et 5.B.4 illustrent les proportions d'analogies de différents niveaux perçues par les élèves lors de la première passation de la tâche et le score associé, ainsi que les nombres moyens d'analogies de chaque niveau perçues lors des trois passations, ainsi que le score total final associé suivant les groupes d'élèves des deux échantillons. Nous nous sommes intéressés de plus aux erreurs commises par les élèves des différents groupes. Nous distinguons comme type d'erreur la sélection de l'intrus, de la non identification de l'analogue de niveau 1 contenant à la fois des similarités de surface et de surface avec le texte source.

**Fig. et Tab. 5.B.2:** Comparaison des proportions d'analogies perçues par les élèves lors de la 1ere passation, suivant les niveaux scolaires A, B ou C de l'échantillon 2 (à gauche), suivant les groupes Standard vs. HP de l'échantillon 1 (à droite).



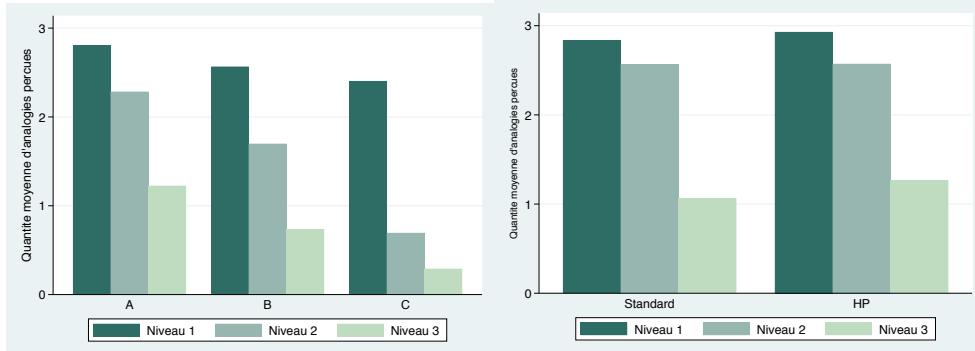
**Fig. et Tab. 5.B.3:** Comparaison des scores des élèves lors de la 1ere passation et de leurs scores totaux sur les trois passations, suivant les niveaux scolaires A, B ou C de l'échantillon 2 (à gauche), suivant les groupes Standard vs. HP de l'échantillon 1 (à droite).



Niveau Scolaire	Elevés N	Proportions au 1er tour % : Moyenne (Err. St)					Score
		Niveau 1	Niveau 2	Niveau 3	Intrus	Erreur1	
A	67	95.52 (3)	73.13 (5)	35.82 (6)	4.47 (3)	4.47 (3)	3.83 (0.3)
B	76	82.89 (4)	53.95 (6)	19.74 (5)	7.90 (3)	13.17 (4)	2.66 (0.2)
C	62	80.65 (5)	19.35 (5)	6.45 (3)	16.12 (5)	9.71 (4)	1.37 (0.2)

Type	Elevés N	Proportions au 1er tour % : Moyenne (Err. St)					Score
		Niveau 1	Niveau 2	Niveau 3	Intrus	Erreur1	
Standard	159	93.08 (2)	82.39 (3)	34.59 (4)	1.89 (1)	4.40 (2)	3.95 (0.2)
HP	42	95.24 (3)	78.57 (6)	35.71 (7)	0 (0)	0 (0)	3.95 (0.3)

**Fig. et Tab. 5.B.4:** Comparaison des quantités d'analogies perçues par les élèves lors des trois passations, suivant les niveaux scolaires A, B ou C de l'échantillon 2 (à gauche), suivant les groupes Standard vs. HP de l'échantillon 1 (à droite).



Niveau Scolaire	Elevés N	Quantité moyenne sur les 3 tours (Err. St)					Score Total
		Niveau 1	Niveau 2	Niveau 3	Intrus	Erreurs1	
A	67	2.81 (0.05)	2.28 (0.10)	1.22 (0.12)	0.21 (0.06)	0.15 (0.05)	12.16 (0.58)
B	76	2.57 (0.08)	1.70 (0.12)	0.74 (0.10)	0.34 (0.07)	0.37 (0.07)	8.74 (0.50)
C	62	2.40 (0.09)	0.69 (0.12)	0.29 (0.10)	0.44 (0.08)	0.35 (0.08)	4.73 (0.40)

Type	Elevés N	Quantité moyenne sur les 3 tours (Err. St)					Score Total
		Niveau 1	Niveau 2	Niveau 3	Intrus	Erreurs1	
Standard	159	2.84 (0.03)	2.57 (0.06)	1.06 (0.08)	0.11 (0.03)	0.12 (0.03)	12.17 (0.40)
HP	42	2.93 (0.05)	2.57 (0.12)	1.26 (0.18)	0 (0)	0.02 (0.02)	13.12 (0.86)

De la même façon que pour la tâche 1, nous avons réalisé des analyses statistiques afin de déterminer les degrés de significativité des différences observées. Les résultats obtenus sont détaillés dans les tableaux 5.B.5 et 5.B.6. Bien que l'ensemble des différences mesurées (à l'exception des taux d'analogie de niveau 2 lors de la 1ere passation chez les HP versus standard) entre les groupes aillent dans le sens de nos hypothèses 2.A et 2.B, seules celles associées à la comparaison des groupes de niveau de l'échantillon 2 sont significatives, celles obtenues pour l'échantillon 1 ne permettent pas de conclure de manière significative pour les comparaisons des groupes HP et Standard, sauf pour ce qui est des erreurs (liées à l'intrus ou à la non sélection de l'anologue 1) qui ne sont jamais commises par les HP.

Comme précédemment, nous avons aussi analysé l'échantillon 1 suivant les groupes de performance scolaire. Les résultats obtenus sont significatifs et alignés avec notre hypothèse 2.B (voir annexes). Nous avons effectué de tests de rapport de vraisemblance comparant l'apport du modèle global intégrant les deux facteurs versus les modèles tenant compte que de l'un ou l'autre des facteurs. Les p values obtenues ne sont pas significatives et ne conduisent à rejeter l'hypothèse nulle que dans le cas du modèle basé sur l'effet Type.

**Tableau 5.B.5:** Tache 2 - Statistiques, effets et p values associés aux réponses des élèves de l'échantillon 2 selon les groupes de niveaux scolaires A, B ou C , lors de la 1ere passation de la tâche.

Statistiques	Niv1	Niv2	Niv3	Intrus	Erreur	Score	TotalScore
<b>Modele parametrique</b>							
<b>Modele</b>	Reg.Logistique	Reg.Logistique	Reg.Logistique	Reg.Logistique	Reg.Logistique	Welch ANOVA	Welch ANOVA
<b>Stat/ pvalue</b>	$z= -2.24$ $p=0.025$	$z = -5.56$ $p=0.000$	$z = -3.81$ $p=0.000$	$z= 2.48,$ $p=0.013$	$z=1.05,$ $p=0.296$	$F(2,193)=19.29,$ $p=0.000$	$F(2,195)=42.42,$ $p=0.000$
<b>Effet (SEM) pvalue</b>							
B vs. A	OR: 0.24 (0.16), $p=0.036$	OR: 0.40 (0.15), $p=0.016$	OR: 0.41 (0.17), $p=0.028$	Non sig	Non sig	-1.17 (0.38), $p=0.008$	-3.43 (0.82), $p=0.002$
C vs. A	OR: 0.18 (0.13), $p=0.019$	OR: 0.08 (0.03), $p=0.000$	OR: 0.11 (0.06), $p=0.000$	OR: 5.64 (4.18), $p=0.020$	Non sig	-2.45 (0.34), $p=0.000$	-7.43 (0.73), $p=0.000$
C vs. B	Non sig	OR: 0.20 (0.6), $p=0.001$	OR: 0.28 (0.17), $p=0.033$	Non sig	Non sig	-1.28 (0.31), $p=0.001$	-4.00 (0.60), $p=0.001$
<b>Modele non-parametrique</b>							
<b>Modele</b>	chi2	chi2	chi2	chi2	chi2	Kruskall Wallis H	Kruskall Wallis H
<b>Stat/ pvalue</b>	$\text{chi2}(2) = 3.21,$ $\text{Fischer } p=0.016$	$\text{chi2}(2) = 68.25$	$\text{chi2}(2) = 37.26,$ $\text{Fischer } p=0.000$	$\text{chi2}(2) = 5.48,$ $\text{Fischer } p=0.084$	$\text{chi2}(2) = 3.21,$ $\text{Fischer } p=0.195$	$\text{chi2}(2)= 44.69,$ $p=0.0001$	$\text{chi2}(2)= 70.02,$ $p=0.0001$
<b>Effet (SEM) pvalue</b>							
B vs. A						$p=0.003$	$p=0.003$
C vs. A						$p=0.000$	$p=0.000$
C vs. B						$p=0.000$	$p=0.000$

**Tableau 5.A.6:** Tache 2 - Statistiques, effets et p values associés aux réponses des élèves de l'échantillon 1selon les groupes Standard et HP lors de la 1ere passation de la tâche.

Statistiques	Niv1	Niv2	Niv3	Intrus	Erreur	Score	TotalScore
<b>Modele parametrique</b>							
<b>Modele</b>	Reg.Logistique	Reg.Logistique	Reg.Logistique	Reg.Logistique	Reg.Logistique	Welch ANOVA	Welch ANOVA
<b>Stat/ pvalue</b>	$z= -0.37$ $p=0.714$	$z = -0.54$ $p=0.592$	$z = -0.02$ $p=0.985$	Predict	Predict	$F(1,192)=0.00,$ $p=0.973$	$F(1,194)=0.42,$ $p=0.519$
<b>Effet (SEM) pvalue</b>	Non sig	Non sig	Non sig			Non sig	Non sig
<b>Modele non-parametrique</b>							
<b>Modele</b>	chi2	chi2	chi2	chi2	chi2	Mann-Whitney U	Mann-Whitney U
<b>Stat/ pvalue</b>	$\text{chi2}(1) = 0.26,$ $\text{Fischer } p=0.460$	$\text{chi2}(1) = 0.3225$	$\text{chi2}(1) = 0.018,$ $p=0.570$	$\text{chi2}(1) = 0.8045,$ $p=0.892$	$\text{chi2}(1) = 1.91,$ $\text{Fischer } p=0.493$	$\text{chi2}(1)= 0.001,$ $p=0.972$	$\text{chi2}(1)= 1.262,$ $p=0.261$
<b>Effet (SEM) pvalue</b>	Non sig	Non sig	Non sig	Non sig	Non sig	Non sig	Non sig

## VI. Discussion

---

En nous appuyant sur les travaux de Chi et al (1981) et des études menées par Smith et al (2013) ainsi que Krieter et al (2016), nous avons élaboré un matériel expérimental basé sur des exercices de catégorisation et de sélection d'analogies s'articulant autour de petits récits de la vie quotidienne dont les dimensions abstraites relèvent du domaine du connu pour des enfants de dix ans. L'usage de ces petits récits nous a permis de comparer cette capacité d'abstraction chez les élèves indépendamment de l'acquisition ou de la consolidation de connaissances spécifiques dans un domaine.

La première tâche constituant le protocole expérimental a permis de mettre en évidence que les élèves avec les meilleurs résultats scolaires ont plus tendance que les autres à baser leurs critères de classification sur des similarités de structure que de surface et d'autre part que cette tendance est aussi caractéristique des enfants à Haut Potentiel. Ces résultats viennent corroborer les hypothèses défendues par Dunbar et Blanchette (2000) ainsi que par Raynal, Clément et Sander (2017) du fondement de l'analogie sur les éléments structurels lorsque les connaissances du sujet lui permettent d'encoder les dimensions abstraites et de négliger les détails de surface.

Les résultats de la deuxième tâche de notre protocole indiquent que les différences observées entre ces élèves de différents niveaux scolaires ne se situent pas seulement au niveau des critères de classification, mais au niveau de la perception même et de l'encodage des éléments communs aux situations. Les données recueillies pour cette deuxième tâche ne nous permettent pas de conclure de manière statistiquement significative pour le groupe HP versus le groupe Standard, bien que les différences mesurées aillent dans le sens de nos hypothèses. Ce qui est significatif en revanche, est l'écart d'erreurs (aussi bien

en ce qui concerne la sélection du texte intrus que de la non identification de l'analogie de niveau 1) par les HP par rapport aux Standard. Il serait très intéressant d'analyser plus en détails les effets de l'interaction du niveau scolaire et de la précocité ainsi que les interactions entre les deux au sein du même échantillon. Pour ce faire, il faudrait disposer d'un plus grand échantillon d'élèves à Haut Potentiel avec des niveaux scolaires différents.

Ces tâches basées sur des petits récits de la vie quotidienne offrent un paradigme original non utilisé auparavant, qui permet de mettre en évidence des points communs et des différences entre élèves dans un domaine décontextualisé, ne requérant aucune connaissance préalable et où ils peuvent être en théorie tous être expert indépendamment de leurs niveaux scolaires. Ces méthodes présentent de plus l'avantage d'être peu couteuses, fiables et faciles à mettre en œuvre dans les salles de classe auprès de larges populations d'enfants. Elles permettent de nombreuses analyses à la fois quantitatives (mais aussi qualitatives) quant aux logiques de classifications et représentations mentales des élèves. Ces jeux de tri de cartes sont d'ailleurs beaucoup utilisés depuis des années en psychologie cognitive, dans des batteries de tests comme par exemple le Wisconsin Card Sorting Test (WCST), et de façon plus récente en ergonomie informatique pour la conception de systèmes d'information (Hannah, 2005).

Les tâches de ce protocole, ainsi que les métriques utilisées pour analyser les résultats ont été élaborées relativement à des catégorisations hypothétiques de surface et de structure préalablement définies. Le grand nombre de classifications ne correspondant à aucune des deux catégorisations hypothétiques, présent dans chaque groupe d'élèves, soulève la question des limitations de la dichotomie habillage de surface/ structure profonde et celle de l'importance des structures induites construites par le sujet en fonction de ses connaissances préalables et des éléments sémantiques de l'énoncé comme support de l'interprétation (Bassok & Olseth, 1995 ; Bassok, 2001). Plusieurs représentations distinctes peuvent donc être associées à une situation et différentes catégorisations être inférées à des niveaux intermédiaires. Afin de tenir compte de ce phénomène, il serait intéressant d'introduire des métriques ne faisant pas référence aux catégorisations hypothétiques, comme cela a été fait dans les travaux de Krieter et al (2016) par le biais de matrices de comparaison

calculées à partir des taux de cooccurrence observées chez différents groupes de sujets. Cette approche repose ainsi sur l'estimation de taux d'appariement à partir des réponses de groupes de sujets prédefinis afin d'évaluer les classifications de nouveaux groupes de sujets.

Par ailleurs, il convient aussi de se poser la question de l'influence de la pragmatique et du contrat didactique dans notre protocole expérimental : Comment les élèves interprètent-ils la consigne ? Dans quelle mesure comprennent-ils le but de l'exercice ? Comprennent-ils l'importance d'une extraction des dimensions abstraites afin de s'en servir comme critère ? Les interactions avec les connaissances naïves et les expériences préalables des élèves ne peuvent pas être négligées. Nous avons été vigilants lors de la passation des expériences dans les classes à ne donner aucune information préalable aux élèves afin d'avoir accès à leurs réponses spontanées, aux éléments qu'ils jugeaient les plus pertinents. Il serait intéressant de considérer une déclinaison « à tris répétés » de la tache 1, comme réalisée dans les études de Smith et de Krieter, dans laquelle après que les élèves aient regroupé les récits de manière libre, il leur serait demandé de les répartir suivant les catégories (SP) prédefinies et ainsi d'observer les différences après que les indications aient été données. Nous pourrions tout à fait imaginer une version de du protocole intégrant une session d'apprentissage, au cours de laquelle seraient expliqués aux élèves les concepts d'éléments de surface et de structure profonde ainsi que le but de la tache au travers d'exemples suivie d'une autre session dédiée à la réalisation des taches.

Certaines limites de notre dispositif expérimental liées à la constitution de notre échantillon HP sont à évoquer. D'une part nous nous sommes basées sur la définition de la douance la plus commune et la plus simple : « est considéré comme HP un enfant dont le QI est supérieur à un certain seuil de 130 » (Feldman, 1982). Si actuellement le QI fait office de mesure adaptée pour déterminer le potentiel intellectuel, de nombreux théoriciens considèrent que ce score ne reflète que partiellement la haute potentialité humaine (Lubart, 2006 ; Sternberg, 1985).

D'autre part, comme c'est le cas pour de nombreuses études impliquant les enfants HP, notre sous-groupe HP n'est pas formé d'enfants choisis aléatoirement dans la population des élèves précoces, mais est formé d'élèves qui ont été diagnostiqués. La détection des enfants précoces n'étant pas systématique en France, cela introduit un biais dans notre échantillon car c'est en effet souvent parce qu'il y a des problèmes qu'un enfant est testé. Or beaucoup d'enfants précoces n'ont pas de problème particulier, ils réussissent bien à l'école et personne ne s'en inquiète (Gauvrit, 2014 ; Gauvrit et Ramus, 2017). Il est donc fort probable que notre groupe d'enfants Standard, constitue des élèves pour lesquels aucun diagnostic HP n'a été posé, contiennent de nombreux enfants HP.

Finalement, les résultats obtenus sont à appréhender en tenant compte de l'origine socio-économique. Il serait intéressant de répliquer cette étude avec un plus grand nombre d'établissements sur lesquels basés notre comparaison des groupes HP versus standard.

## CONCLUSION

De nos jours, de plus en plus de réformes dans le monde de l'éducation portent sur le fait d'apprendre aux élèves à apprendre et à penser, à faire preuve d'esprit critique. L'aspiration de l'école ne vise plus seulement l'acquisition de connaissances (qui sont aujourd'hui accessibles à tous de plus en plus facilement par le biais d'internet et des réseaux sociaux) mais l'organisation experte de ces connaissances en créant des connexions entre concepts, la flexibilité mentale de se mouvoir d'une catégorisation à une autre, la capacité d'aborder une situation sous plusieurs angles afin d'en saisir le plus pertinent (Scheibling-Sève, Eichi, Sander et Pasquinelli, 2016).

Pour mesurer les progrès réalisés dans ce domaine, il y a besoin d'outils pédagogiques adaptés. Les programmes actuels appellent à des tâches interdisciplinaires connectant les connaissances acquises dans différents domaines et favorisant le transfert interdisciplinaire des connaissances acquises à l'école des situations de la vie quotidienne.

En effet, si les professeurs disposent de nombreux outils pour évaluer la compréhension d'un texte, la résolution d'un problème ou l'exactitude des connaissances acquises par les élèves, très peu permettent d'évaluer la façon dont ils organisent leurs connaissances. Capacité, que la recherche s'accorde à désigner comme étant un élément clé du développement de l'expertise dans un domaine et corrélé à la capacité de résolution de problèmes. Les premiers travaux ayant mis en évidence les différences d'organisation des connaissances entre novices et experts viennent de la recherche en sciences de l'éducation en physique. Les travaux pionniers de Chi et al (1981) ont ensuite inspirées de nombreuses recherches dans différents domaines. Ces recherches (telles celles de Smith et al (2013) en biologie et de Krieter et al (2016) en chimie) portent sur l'élaboration d'outils pédagogiques visant non plus à évaluer l'acquisition d'un concept scientifique précis, ni l'exactitude d'une stratégie de résolution de problèmes, mais à comparer l'organisation conceptuelle des connaissances avec celles des experts dans ces domaines. Il n'existe pas encore cependant d'outils méthodologiques permettant d'étudier les représentations mentales des élèves, leurs capacités à percevoir des analogies ou à faire preuve de flexibilité cognitive.

Notre étude ouvre donc la porte à des travaux ultérieurs à la fois en termes de déclinaison de ces outils, et de leurs inclusions dans des études longitudinales ainsi que dans différents curriculums scolaires. Ces exercices peuvent en effet aussi bien servir d'outils d'évaluation que d'exercices d'apprentissage visant à développer la flexibilité cognitive ainsi qu'une certaine conscience métacognitive des élèves. Car adopter une nouvelle représentation nécessite de faire preuve d'une certaine distance vis-à-vis d'une représentation spontanée. Ainsi une méthode d'apprentissage basée sur l'encodage sémantique, dans la lignée des travaux de Gamo, Nogry et Sander (2011) dans le cadre de la résolution de problèmes arithmétiques, mais s'articulant ici autour de ces petits récits de la vie quotidienne pourrait être envisagée afin de développer leur capacité d'abstraction et à changer de point de vue.

Ces considérations amènent à s'interroger sur le concept d'intelligence. La définition de l'intelligence ainsi que la question d'une faculté d'intelligence générale ont fait l'objet de nombreuses controverses. En 1986, plus d'une

vingtaine d'experts en psychologie ont été interrogés pour donner une définition de l'intelligence, mais aucun consensus ne s'est dégagé (Sternberg, 1986). Les études scientifiques les plus récentes définissent l'intelligence générale comme « une propriété émergente, résultant de la cascade de facteurs environnementaux, génétiques, cérébraux et cognitifs qui influent sur la performance aux différents tests » (Ramus, 2012).

Dans leur ouvrage, Hofstadter et Sander (2013) posent la question de savoir « si l'intelligence ne réside-t-elle pas au moins en partie dans la quantité de concepts que l'on possède et dans la finesse des réseaux qui lient ces concepts » et définissent l'intelligence comme « l'art d'aller droit au but, au cœur des choses, à l'essentiel rapidement et de manière fiable ». En effet, la sélection de la catégorisation optimale pour saisir l'essence tout en conservant la subtilité de la situation pourrait caractériser une intelligence des choses. Car l'abstraction, qui permet la construction de la catégorie et de son extension peut cependant aussi appauvrir la représentation dès lors qu'elle est en excès. Les êtres humains s'éloignent toujours spontanément du niveau littéral, la difficulté réside dans la sélection du niveau d'abstraction pertinent, particulièrement s'il est profond. La flexibilité cognitive, qui décrit la capacité à faire varier les dimensions selon lesquelles se fait l'encodage et la catégorisation d'une situation fait partie intégrante selon eux de cette définition.

Si le rôle essentiel de l'analogie semble établi dans les activités de compréhension, de résolution de problèmes ainsi que dans l'acquisition de connaissances et que de nombreuses théories ont été élaborées, il serait intéressant de s'interroger sur sa place dans les théories de l'intelligence. En effet l'analogie, qui constitue une composante essentielle de la théorie générale de l'intelligence de Spearman (1923) y est considérée sous la forme de tâches d'analogie classique du type: « A est B, ce que C'est à ... », présentes dans la plupart des tests d'évaluation de l'intelligence (via par ex. les matrices de Raven ou la tâche des dominos). Ces tâches d'analogie classiques ont aussi été utilisées par Piaget pour étudier le développement cognitif des enfants (Piaget et al, 1977). Si les évolutions de la conception de l'intelligence ont amené la prise en compte en considération de la créativité comme dimension de l'intelligence, comme dans la théorie triarchique de l'intelligence (Sternberg, 1985), la

description étude de l'analogie semble rester relativement circonscrite au facteur componentiel de l'intelligence, en lien avec le QI.

De plus il est intéressant de considérer la position défendue par Stanovich (2014; 2016) et de ceux qui insistent pour différencier les notions d'esprit critique et d'intelligence telle que mesurée par le QI. Il est vrai que les tests de QI n'ont pas été élaborés pour mesurer l'esprit critique au sens de la définition de Halpern (1988), et ont plus tendance à tester des précurseurs des compétences nécessaires à l'esprit critique. La compréhension d'un texte, d'une situation, la perception d'analogies classiques sont des prérequis pour percevoir la structure profonde d'une situation. Pour Stanovich, l'esprit critique est fortement lié à la capacité d'identifier les biais classiques de raisonnement. Les expériences de Kahneman (2012) montrent qu'une haute intelligence ne protège pas des erreurs commises par des heuristiques trop rapides. Mais comment alors concilier ce concept de pensée rationnelle à la définition de l'intelligence telle que présentée par Ramus (2012), comment tenir compte les différents aspects de l'analogie et de la perception de l' « essence » d'une situation ? Nombreuses questions, que nous souhaiterions continuer d'étudier...

## VII. Bibliographie

---

- American Association for the Advancement of Science. (2011). Vision and Change in Undergraduate Biology Education: A Call to Action, Washington, DC.
- Akin O. (1982). The psychology of architectural design, Londres, Pro.
- Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory and Cognition*, 10, 82–93.
- Barsalou, L. W. (1983). Ad hoc categories. *Memory and Cognition*, 11, 211–227.
- Barsalou, L. W. (1991). Deriving categories to achieve goals. In G. H. Bower (Ed.), *The psychology of learning and motivation* : Vol. 27 (pp. 1–64). New York : Academic Press.
- Bassok, M. (2001). Semantic alignments in mathematical word problems. *The analogical mind: Perspectives from cognitive science*, 401-433.
- Blanchette, I., Dunbar, K. (2000). How analogies are generated: The roles of structural and superficial similarity. *Memory and Cognition*, vol 28, no. 1.
- Boster J. S., Johnson J. C. (1989). Form or function : A comparison of expert and novice judgments of similarity among fish, *American Anthropologist*, 91, 866-888.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How People Learn Brain, Mind, Experience, and School*. Washington DC National Academy Press.
- Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. New York : Wiley.
- Chatard-Pannetier, A., Brauer, M., Chambres, P., et al. (2002). Représentation, catégorisation et évaluation : différences entre experts et novices dans le domaine des meubles d'antiquité. *L'Année Psychologique*, 102, 423–448.
- Chi, M. T. H., Feltovitch, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121–152.
- Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *J Verbal Learning Verbal Behav.*, 8, 240–248
- Dubois S., Mohib N., Oge D., Schenk E., Sonntag M. (2005). Connaissances et reconnaissance de l'expert. *Les cahiers de l'INSA de Strasbourg*, pp.89-108.
- Dunbar, K. & Blanchette, I. (2001). The invivo/invitro approach to cognition: the case of analogy. *Trends in Cognitive Sciences*, 5, 334-339.
- Feldman D. (1982). A developmental framework for research with gifted children, New directions for child development : developmental approaches to giftedness and creativity, XVII, p. 31-45.
- Gamo, S., Sander, E. & Richard, JF. (2010). Transfer of strategies by semantic recoding in arithmetic problem solving. *Learning and Instruction*, 20, 400-410.
- Gauvrit, N. (2014). *Les Surdoués ordinaires*, Presses Universitaires de France.

- Gentner, D. (1983). Structure-mapping: a theoretical framework for analogy-making. *Cognitive Science*, 7(2), 155-70.
- Gentner, D. (1989). Mechanisms of analogical learning. In S. Vosniadou & A. Ortony, (Eds.). *Similarity and analogical reasoning*, 199-241. London: Cambridge University Press.
- Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology*, 393-408.
- Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology* 12, 306-355.
- Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15(1), 1-38.
- Gough, P. B., & Tunmer, W. E. (1986). Decoding, Reading, and Reading Disability. *Remedial and Special Education*, 7, 6-10.
- Halpern, D. F. (2002). *Thought and knowledge: An introduction to critical thinking*. Routledge.
- Halpern, D. F. (2007). *Halpern critical thinking assessment using everyday situations: Background and scoring standards*. Claremont, CA: Claremont McKenna College.
- Hannah, S. (2005). Sorting out card sorting: Comparing methods for information architects, usability specialists, and other practitioners. MS thesis. Portland, OR: University of Oregon, Applied Information Management Department.
- Hampton, J. A. (1979). Polymorphous concepts in semantic memory. *Verbal Learning Verbal Behav.*, 18, 441–461.
- Hofstadter, D., & Sander, E. (2013). *L'Analogie : au cœur de la pensée*. Paris : Odile Jacob.
- Holyoak, K. J. (1985). The pragmatics of analogical transfer. In G. H. Bower (Eds.), *The psychology of learning and motivation*, vol. 19, 59-87. New York, Academic Press.
- Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition*, 15(4), 332-340.
- Hull, C. L. (1920). Quantitative aspects of the evolution of concepts. *Psychological Monographs*, XXVIII(1.123), 1–86.
- Johnson, K. E., & Mervis, C. B. (1997). Effects of varying levels of expertise on the basic level of categorization. *J Exp Psychol : General*, 126, 248–277.
- Kahneman D. (2011). *Thinking, Fast and Slow* , Farrar, Straus and Giroux.
- Keil, F.C. (1989). Concepts, kinds and cognitive development. Cambridge, MA: MIT Press.
- Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review*, 92, 109-129.
- Kintsch, W. (1998). *Compréhension: a paradigm for cognition*. New York: Cambridge University Press.
- Krieter F.E., Julius R.W., Tanner K.D., Bush S.D., and Scott G.E. (2016). "Thinking Like a Chemist: Development of a Chemistry Card-Sorting Task To Probe Conceptual Expertise," *Journal of Chemical Education*, v.93, p. 811.

- Lubart T.I. (Dir.) (2006). Enfants exceptionnels. Précocité intellectuelle, haut potentiel et talent. Paris : Bréal.
- McCloskey, M., & Glucksberg, S. (1978). Natural categories : Well defined or fuzzy set ? Memory et Cognition., 6, 462–472.
- Medin, D. L., & Ross, B. H. (1989). The specific character of abstract thought: Categorization, problem-solving, and induction. New York: Cambridge Univ. Press.
- Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. Psychological Review, 85(3), 207-238.
- Mervis, C. B., Catlin, J., & Rosh, E. (1976). Relationships among goodness of example, category norms, and word frequency. Bulletin of the Psychonomic Society, 7, 283–294.
- Morris, M. W., & Murphy, G. L. (1990). Converging operations on a basic level in event taxonomies. Memory et Cognition., 18, 407–418.
- McKoon, G. & Ratcliff, R. (1992). Pronoun resolution and discourse models. Journal of Experimental Psychology: Learning, Memory, and Cognition.
- Murphy G.L. (2004) The Big Book of concepts. Cambridge, MA : The MIT Press.
- Murphy, G. & Medin, D. (1985), The Role of Theories in Conceptual Coherence, Psychological Review 92, pp. 289–316.
- Naglieri J., Kaufman J. (2001), Understanding intelligence, giftedness and creativity using PASS theory, Roeper review, XXIII, 3, p. 151-156.
- Nathan, M. J., Kintsch, W., & Young, E. (1992). A theory of algebra-word-problem comprehension and its implications for the design of learning environments. Cognition and Instruction, 9(4), 329–389.
- Nosofsky RM. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. Journal of Experimental Psychology: Learning, Memory, & Cognition. 14:700–708.
- Novick, L. (1988). Analogical transfer, problem similarity, and expertise. Journal of Experimental Psychology: Learning, Memory, and Cognition, 14, 510-520.
- Novick, L. R. & Holyoak, K. J. (1991). Mathematical problem solving by analogy. Journal of Experimental Psychology: Learning, Memory, and Cognition. Vol 17 (3). 398-415
- Pasquinelli E., Zimmerman G., Farina M. (2018). Esprit scientifique, Esprit critique, Editions Le Pommier.
- Ramus, F. (2012). L'intelligence dans tous ses états. L'intelligence. Cerveau Psycho.
- Ramus, F., & Gauvrit, N. (2017). La légende noire des surdoués. La Recherche, Mars 2017.
- Raynal, L; Clément, E., & Sander, E. (2017). Challenging the superficial similarity account for analogical. Congrès de la Société Française de Psychologie.
- Renzulli J. (1986). The three ring conception of giftedness : a developmental model for creative productivity, dans Sternberg (R. J.), Davidson (J. E.), Conception of giftedness, New York, Cambridge university press, p. 53-92.

Reusser, K. (1990). From text to situation to équation: Cognitive simulation of understanding and solving mathematical word problems. In H. Mandl, E. De Corte, S. N. Bennett & H. F. Friedrich (Eds.), Learning & Instruction: European Research in an International Context (Vol. 2, pp. 477-498). Oxford: Pergamon Press.

Ripoll T. (1992). La recherche sur le raisonnement par analogie : objectifs, difficultés et solutions. In: L'année psychologique. vol. 92, n°2. pp. 263-288.

Ripoll T., Coulon D. Le raisonnement par analogie : une analyse descriptive et critique des modèles du mapping. In: L'année psychologique. 2001 vol. 101, n°2. pp. 289-323.

Rosch, E. (1975). Cognitive representations of semantic categories. *J Exp Psychol : General*, 104, 192–233.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and Categorization* (pp. 27–48). Hillsdale, NJ : Erlbaum.

Rosch, E., & Mervis, C. B. (1975). Family resemblances : studies in the internal structure of categories. *Cognitive Psychology*, 7, 573–605.

Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(4), 629.

Rumelhart, D. E., & Ortony, A. (1977). The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro & W. E. Montague (Eds.).

Sander, E. (2000). L'analogie, du Naïf au Crétif : analogie et catégorisation. L'Harmattan : Paris.

Sander, E. (2016). Le développement conceptuel. *Psychologie du développement*. Elsevier, p107-117.

Schank R.C.,& Abelson R.P. (1977). Scripts, Plans, Goals and Understanding: an Inquiry into Human Knowledge Structures.

Scheibling-Sève, C., Eichi, A., Sander, E., Pasquinelli, E. (2016). Méthode d'apprentissage et évaluation en résolution de problèmes d'arithmétiques : outils progressifs pour développer la flexibilité des stratégies de résolution, Colloque International « Evaluation en mathématiques : dispositifs, validités & pratiques », ESPE de l'Académie de Créteil, 21-22 novembre 2016.

Smith J.I., Combs E.D., Nagami P.H., Alto V.M., Goh H.G., Gourdet Maa, Hough C.M., Nickell A.E., Peer A.G., Coley J.D., et al. (2013). Development of the Biology Card Sorting Task to measure conceptual expertise in biology. *CBE Life Sci Educ* 12, 628–644.

Smith, E. E., & Medin, D. L. (1981). Categories and concepts. Cambridge : Harvard University Press.

Spearman, C. (1923). The nature of intelligence and the principles of cognition. Un vol. in-8°, 358 pp.

Stanovich, K. E. (2016). The Comprehensive Assessment of Rational Thinking. *Educational Psychologist*, 51, 23-34.

Sternberg R. (1985). Beyond IQ : a triarchic theory of human intelligence, New York, Cambridge university press.

- Sternberg R., Lubart T. (1992). Creative giftedness in children, dans Tannenbaum (A. J.), Klein (P. S.), To be young and gifted, Westport, Ablex publishing, p. 33-51.
- Treffinger D. (1980). The progress and peril of identifying creative talent among gifted and talented students, Journal of creative behavior, XIV, 1, p. 20-34.
- Trench, M., & Minervino, R. A. (2015). The role of surface similarity in analogical retrieval: Bridging the gap between the naturalistic and the experimental traditions. Cognitive science, 39(6), 1292-1319.
- Van Dijk, T. A. (1977). Semantic Macro-Structures and Knowledge Frames in Discourse Compréhension. In M. A. Just & P . A. Carpenter (Eds.), Cognitive Processes in Compréhension (pp. 3-32). Hillsdale: Lawrence Erlbaum associâtes,
- Van Dijk, T. A & Kintsch, W. (1983). Strategies of discourse comprehension. New York Press.
- Willingham, D. T. (2007). Critical Thinking Why Is It So Hard to Teach American Educator, 31, 8-19.
- Wisniewski, E. J., & Medin, D. L. (1994). On the interaction of theory and data in concept learning. Cognitive Science, 18(2), 221-281.
- Wittgenstein, L. (1953). Philosophical investigations. Oxford : Basil Blackwell.