



<http://portaildoc.univ-lyon1.fr>

Creative commons : Paternité - Pas d'Utilisation Commerciale -  
Pas de Modification 2.0 France (CC BY-NC-ND 2.0)



<http://creativecommons.org/licenses/by-nc-nd/2.0/fr>

**N° de mémoire 2139**

Mémoire d'Orthophonie

présenté pour l'obtention du

**Certificat de capacité d'orthophoniste**

Par

**BETIN Elodie**

**Développement de la compréhension de la valeur positionnelle :**

**Étude de la validité du module de numération de la batterie**

**d'évaluation Examath 5-8**

Directrice de Mémoire

**LAFAY Anne**

Année académique

**2020-2021**

Directeur ISTR

**Xavier PERROT**

**Equipe de direction du département d'orthophonie :**

Directeur de la formation

**Agnès BO**

Coordinateur de cycle 1

**Claire GENTIL**

Coordinateur de cycle 2

**Solveig CHAPUIS**

Responsables de l'enseignement clinique

**Claire GENTIL**

**Ségolène CHOPARD**

**Johanne BOUQUAND**

Responsables des travaux de recherche

**Mélanie CANAULT**

**Floriane DELPHIN-COMBE**

**Claire GENTIL**

Responsable de la formation continue

**Johanne BOUQUAND**

Responsable du pôle scolarité

**Rachel BOUTARD**

Secrétariat de scolarité

**Anaïs BARTEVIAN**

**Constance DOREAU KNINDICK**

**Céline MOULART**

## **1. UNIVERSITE CLAUDE BERNARD LYON 1**

Président  
**Pr. FLEURY Frédéric**

Vice-président CFVU  
**Pr. CHEVALIER Philippe**

Vice-président CA  
**Pr. REVEL Didier**

Vice-président CS  
**M. VALLEE Fabrice**

Directeur Général des  
Services  
**M. VERHAEGHE Damien**

### **1.1 Secteur Santé**

U.F.R. de Médecine Lyon Est  
Doyen **Pr. RODE Gilles**

Institut des Sciences Pharmaceutiques et  
Biologiques  
Directrice **Pr. VINCIGUERRA Christine**

U.F.R de Médecine et de maïeutique  
- Lyon-Sud Charles Mérieux  
Doyenne **Pr. BURILLON Carole**

Institut des Sciences et Techniques  
de la Réadaptation (I.S.T.R.)  
Directeur **Dr. PERROT Xavier**

U.F.R d'Odontologie  
Directrice **Pr. SEUX Dominique**

Département de Formation et Centre  
de Recherche en Biologie Humaine  
Directrice **Pr. SCHOTT Anne-Marie**

### **1.2 Secteur Sciences et Technologie**

U.F.R. Faculté des Sciences  
Administrateur provisoire  
**M. ANDRIOLETTI Bruno**

Institut des Sciences Financières et  
d'Assurance (I.S.F.A.)  
Directeur **M. LEBOISNE Nicolas**

U.F.R. Biosciences  
Administratrice provisoire  
**Mme GIESELER Kathrin**

Observatoire Astronomique de Lyon  
Directeur **Mme DANIEL Isabelle**

U.F.R. de Sciences et Techniques  
des Activités Physiques et Sportives  
(S.T.A.P.S.)  
Directeur **M. VANPOULLE Yannick**

Ecole Supérieure du Professorat et  
de l'Éducation (E.S.P.E.)  
Administrateur provisoire  
**M. Pierre CHAREYRON**

Institut National Supérieure du  
Professorat et de l'Éducation  
(INSPé)  
Directeur **M. CHAREYRON Pierre**

POLYTECH LYON  
Directeur **M. PERRIN Emmanuel**

Institut Universitaire de Technologie  
de Lyon 1 (I.U.T.LYON 1)  
Directeur **M. VITON Christophe**

## Résumé

La compréhension du concept de la valeur positionnelle est essentielle au traitement des nombres. Sa maîtrise est un prédicteur des compétences arithmétiques futures. C'est une compétence difficile à acquérir, également pour les enfants aux développements typiques. Évaluer la compréhension approfondie de la valeur positionnelle est primordial ; pourtant, peu d'outils à la disposition des orthophonistes permettent d'évaluer cette compétence pour les enfants en apprentissage, c'est-à-dire entre 5 et 8 ans. Les objectifs de cette recherche étaient de vérifier la validité concomitante, la validité de construit en lien avec les caractéristiques des individus (la classe) et la cohérence interne du module de numération de la nouvelle batterie Examath 5-8. L'échantillon était composé de 15 enfants scolarisés en GSM, 12 enfants scolarisés en CP et 14 enfants scolarisés en CE1, sans suivi orthophonique pour des difficultés mathématiques, dans deux écoles différentes. Le matériel utilisé était composé des épreuves de numération des batteries d'Examath 5-8 et Tedi-Math ainsi que le PicPVT. Les résultats ont montré que le module de numération de la batterie d'évaluation Examath 5-8 possède une bonne validité concomitante, une bonne validité de construit en lien avec les caractéristiques des individus (la classe) et une bonne cohérence interne. L'établissement d'une norme critériée a permis de montrer que les nouvelles tâches d'évaluation de la valeur positionnelle sont adaptées et permettent d'observer les compétences en numération des enfants en apprentissage. L'analyse des productions des enfants a permis de confirmer les modèles théoriques actuels et d'apporter quelques suggestions. Ces éléments suggèrent une bonne validité de contenu. D'autres propriétés psychométriques doivent être évaluées pour permettre la validation de cette nouvelle batterie, comme la validité de surface, la validité prédictive ou la sensibilité. Ce nouvel outil sera une grande avancée pour la clinique orthophonique.

### Mots clés

Mathématiques – Évaluation – Numération – Valeur positionnelle – Enfant

## **Abstract**

Understanding the concept of place value is essential to number processing. Its command is a predictor of future arithmetic skills. It is a difficult skill to acquire, even for children with typical development. Therefore, assessing the thorough understanding of place value is crucial. However, there are few tools available to speech-language therapists to assess this skill in learners between the ages of five and eight years. The aim of this research was to verify (1) the internal consistency of the Examath 5-8 battery numeracy module, (2) the concurrent validity and (3) the construct validity in relation to individual characteristics (grade). The sample was composed of 15 children in kindergarten, 12 children in first grade, and 14 children in second grade. These children were from two different schools and had no intervention from a speech therapist of difficulties in mathematics. Numeracy tasks from Examath 5-8, Tedi-Math, and PicPVT were used. The results showed that the numeracy module of the Examath 5-8 demonstrated good internal consistency, good concurrent validity, and good construct validity in relation to individual characteristics (grade). The creation of a criterized norm allowed to demonstrate that the new place-value assessment tasks are adapted and allow to observe the numeracy skills of learning children. The analysis of the children's performances confirmed recent theoretical models and provided certain suggestions. These elements suggest good content validity. In order to develop the new test battery Examath 5-8, other psychometric aspects need to be investigated, such as surface validity, predictive validity, or sensibility. This new tool will be a major advancement for the speech-and-language clinical practice.

### **Key words**

Mathematics – Evaluation – Numeration – Place value – Child

## Remerciements

Je tiens à remercier ma directrice de mémoire Anne Lafay, pour son immense disponibilité et son soutien sans faille durant ces deux dernières années. Merci de m'avoir offert cet accompagnement, pour mes premiers pas dans la recherche.

Je remercie également les deux directrices d'école qui ont cru en mon projet, qui ont offert de m'accueillir malgré les contraintes liées à la pandémie, et sans qui ce mémoire n'aurait pu avoir lieu.

Je remercie chaleureusement tous les enfants qui se sont porté-es volontaires pour participer à ce projet. Je les remercie pour leurs implications sérieuses, leurs questions lumineuses et leur bonne humeur.

Enfin je suis très reconnaissante envers mes proches, tous endurants qu'il-elle-s sont dans leur soutien.

## Sommaire

<b>I. PARTIE THEORIQUE .....</b>	<b>1</b>
1 CONTEXTE.....	1
2 DEFINITION DE LA VALEUR POSITIONNELLE .....	1
3 LA COMPREHENSION DE LA VALEUR POSITIONNELLE COMME PREDICTEUR.....	2
4 INFLUENCE DE LA LINGUISTIQUE SUR L'APPRENTISSAGE DE LA VALEUR POSITIONNELLE ..	2
5 LE DEVELOPPEMENT DE LA VALEUR POSITIONNELLE.....	3
5.1. <i>Avant l'apprentissage formel</i> .....	3
5.2. <i>L'apprentissage formel</i> .....	4
6 ÉTAT DES LIEUX DES OUTILS D'ÉVALUANT LA VALEUR POSITIONNELLE .....	6
6.1. <i>Dans la recherche</i> .....	6
6.2. <i>En clinique</i> .....	7
7 PROPRIÉTÉS PSYCHOMÉTRIQUES D'UN OUTIL D'ÉVALUATION.....	8
7.1. <i>Validité</i> .....	8
7.2. <i>Fidélité</i> .....	9
7.3. <i>Définition d'une norme critériée</i> .....	9
8 PROBLÉMATIQUES.....	9
9 OBJECTIFS DE RECHERCHE .....	10
<b>II. METHODE .....</b>	<b>11</b>
1 PARTICIPANTS .....	11
2 MATÉRIELS .....	11
2.1. <i>Examath 5-8</i> .....	11
2.1.1 Évaluation de la maîtrise des codes .....	11
2.1.2 Évaluation de la compréhension du système de numération .....	12
2.2. <i>Tedi-math</i> .....	13
2.3. <i>PicPVT</i> .....	14
3 PROCÉDURE GÉNÉRALE .....	14
<b>III. RESULTATS .....</b>	<b>15</b>
1 HYPOTHÈSE 1 – COHÉRENCE INTERNE .....	15
2 HYPOTHÈSE 2 – VALIDITÉ CONCOMITANTE .....	15
2.1. <i>Épreuve de Transcodage en Lecture</i> .....	15



2.2. Épreuve de Transcodage en Dictée .....	15
2.3. Épreuve Valeur des chiffres.....	15
2.4. Épreuve UDC .....	16
3 HYPOTHESE 3 – VALIDITE DE CONSTRUIT EN LIEN AVEC LES CARACTERISTIQUES DES INDIVIDUS : LE NIVEAU SCOLAIRE.....	16
3.1. Épreuve de Transcodage en Lecture .....	16
3.2. Épreuve de Transcodage en Dictée .....	18
3.3. Épreuve UDC .....	19
3.4. Épreuve Valeur des chiffres.....	19
3.5. Épreuve Décomposition d'addition .....	19
4 HYPOTHESE 4 – NORME CRITERIEE .....	19
5 ANALYSE QUALITATIVE DES ERREURS .....	20
5.1. Erreurs de transcodage.....	21
5.2. Erreurs de valeur positionnelle .....	21
<b>IV. DISCUSSION .....</b>	<b>23</b>
1 INTERPRETATIONS DES RESULTATS.....	23
1.1. Hypothèse 1 – Cohérence interne du module de numération .....	23
1.2. Hypothèse 2 – Validité concomitante du module de numération .....	23
1.3. Hypothèse 3 – Validité de construit .....	23
1.4. Hypothèse 4 – Établissement d'une norme critériée et analyse qualitative des erreurs.....	25
2 LIMITES DE L'ETUDE ET PERSPECTIVES DE RECHERCHE .....	28
3 IMPLICATIONS CLINIQUES.....	29
<b>V. CONCLUSION .....</b>	<b>30</b>
<b>VI. BIBLIOGRAPHIE .....</b>	<b>31</b>
<b>VII. ANNEXES .....</b>	<b>38</b>

## I. Partie théorique

### 1 Contexte

Le système numérique indo-arabe utilisé actuellement est le fruit de longues évolutions culturelles (Dehaene, 2010). Son efficacité réside dans sa capacité à coder des nombres infiniment grands avec un petit ensemble fini de symboles, grâce à son système de position (Herzog et al., 2019). La compréhension et l'utilisation de ce système de numération est centrale à toute activité numérique (Nuerk et al., 2015), et plus particulièrement, une bonne compréhension du concept de la valeur positionnelle est essentielle à l'apprentissage des mathématiques (Chan & Ho, 2010 ; Chan et al., 2014). Cependant, la compréhension du concept de la valeur positionnelle semble être un défi à relever pour les enfants au développement typique : des difficultés en lien avec ce concept ont été signalées à presque tous les niveaux scolaires, dans différents pays et pour des apprenants de tout niveau de performances (Fritz & Ricken, 2008 ; Fuson, Wearne, et al., 1997 ; Gervasoni, 2007 ; MacDonald, 2008). Ces difficultés sont relevées dans divers études mais la recherche ne fait pas état de sa prévalence (Chan & Ho, 2010 ; Desoete, 2015 ; Fritz & Ricken, 2008 ; Kamii, 1986; Nuerk et al., 2015 ; Ross, 1989). Parmi les enfants manipulant les nombres, il est observé que beaucoup d'entre eux n'acquièrent pas une connaissance approfondie de la valeur positionnelle (Fritz & Ricken, 2008 ; Fuson, Wearne, et al., 1997; Kamii, 1986 ; Ross, 1989). Or, seule une compréhension approfondie peut permettre de manipuler les nombres lors de calculs simples ou complexes (Cawley et al., 2007). L'objectif de ce mémoire est de vérifier la validité d'un nouvel outil d'évaluation de la compréhension de la VP. La partie suivante exposera l'état des connaissances actuelles sur le développement de la compréhension de la VP, les outils à dispositions des orthophonistes pour la mesurer et les caractéristiques psychométriques nécessaires à un outil d'évaluation. Ensuite la partie expérimentale sera exposée : la méthode et les résultats seront détaillés. Enfin, ce travail sera discuté en ultime partie.

### 2 Définition de la valeur positionnelle

Le système de numération en notation indo-arabe repose sur trois valeurs (Chandler & Kamii, 2009 ; Fuson, Smith, et al., 1997). Concernant la valeur de position, la position des chiffres dans le nombre code une valeur particulière, ainsi l'unité se trouve à droite, la dizaine à sa gauche, la centaine à gauche de la dizaine, etc. D'après la valeur multiplicative, chaque chiffre indique le nombre de fois que la valeur est

présente dans le nombre. Enfin, la valeur somme indique que toutes les valeurs des chiffres sont additionnées entre elle pour former la valeur du nombre. De ces trois valeurs, l'acquisition de la valeur positionnelle (VP) est l'objet de nombreuses difficultés chez l'enfant (Hanich et al., 2001).

### **3 La compréhension de la valeur positionnelle comme prédicteur**

La compréhension du système de numération à valeur positionnelle est primordiale pour le traitement des nombres à plusieurs chiffres (Nuerk et al., 2015) et essentielle à la réalisation de nombreuses tâches numériques et arithmétiques. Ainsi Laski et al., (2016) ont montré qu'une compréhension de la VP dès la maternelle soutient de meilleurs choix de stratégies de calcul en classe de CE1 (cours élémentaire première année), c'est-à-dire une utilisation de la décomposition de nombres plutôt que le comptage ou la récupération en mémoire de faits arithmétiques. Cawley et al. (2007) interprètent que les enfants qui manipulent les nombres principalement par le comptage ou la récupération en mémoire de faits arithmétiques n'ont pas acquis une compréhension de la VP approfondie. Moeller et al. (2011) ont démontré que la réussite à la tâche de comparaison de nombres en classe de CP (cours préparatoire) est un marqueur de la compréhension de la VP et que ce marqueur est un prédicteur des performances mathématiques en CE2 (cours élémentaire deuxième année). Réciproquement, de faibles compétences de la VP peuvent être prédictrices - et une source - de difficultés en mathématiques (Cawley et al., 2007 ; Desoete, 2015). Chan et al. (2014) ont également montré une prédiction des performances mathématiques en CE1 par une tâche évaluant la compréhension de la VP. Ils ont décrit trois types de trajectoires développementales liées à cette compétence : les enfants déjà performants qui ont peu de marges de progression, les enfants moins performants qui rattraperont les compétences du premier groupe avec une belle progression et les enfants en difficultés qui maintiendront leur écart en CE1. Ainsi, la mesure de la compréhension de la VP présente un enjeu de taille pour l'évaluation des troubles mathématiques.

### **4 Influence de la linguistique sur l'apprentissage de la valeur positionnelle**

Le système indo-arabe est adopté uniformément dans le monde mais le code oral qui s'y rattache est dépendant du système linguistique de chaque langue. Ainsi, les mots-nombres sont plus ou moins transparents selon les langues et les variations concernent principalement la construction morphologique des nombres à plusieurs

chiffres. Par exemple, en chinois le nombre « onze » se dit « 1-10 et 1 » : la construction est transparente, alors qu'en français il s'agit d'un mot nouveau, au même titre que « sept » ou « neuf » et aucune information n'est donnée sur la structuration en dizaines et unités. En allemand et en néerlandais, le mot-nombre est construit avec l'ordre « unité-dizaine » (27 est dit « sept-et-vingt »), alors que l'italien et l'anglais sont plus transparents et construisent le mot-nombre comme le code indo-arabe (c'est-à-dire « dizaine-unité »). Pixner et al. (2011), mais aussi Dowker et Roberts (2015), ont étudié ces différences et ont conclu qu'un système de mots-nombres transparent est facilitateur de l'apprentissage de la VP. Mark et Dowker (2015) nuancent et observent que le système linguistique a une faible influence sur l'apprentissage de la VP. En conclusion, le système linguistique influence l'apprentissage de la VP mais n'est pas l'unique facteur.

## **5 Le développement de la valeur positionnelle**

### **5.1. Avant l'apprentissage formel**

De nombreuses études montrent qu'avant l'apprentissage explicite de la VP, les enfants au contact des nombres ont déjà extrait quelques règles. McGuire et Kinzie (2013) ont noté que les enfants en moyenne section de maternelle ont plus de facilités à isoler l'unité d'un nombre (ex : 3 dans 23) que la dizaine (ex : 20 dans 23). Ceci s'explique par le fait que pour envisager une dizaine, il faut concevoir le concept de regroupement, c'est-à-dire considérer un groupe comme une unité. Selon Gelman et Meck (1983), le concept de l'unité est acquis autour de 3 ans. Le concept de regroupement, allant à l'encontre du concept de l'unité, est acquis plus tardivement (Fosnot & Dolk, 2001). Ce concept de regroupement est essentiel à l'acquisition de la VP (Cobb & Wheatley, 1988). Byrge et al. (2014) ont observé les compétences des enfants en moyenne section (MS) et grande section de maternelle (GSM) sur une tâche de dictée de nombres. Ils ont observé une erreur fréquente : « l'écriture littérale » (aussi traduite écriture terme à terme ou élargie), c'est-à-dire une notation où les chiffres sont dans le bon ordre, mais avec l'ajout de zéros supplémentaires (par exemple, écrire « 600402 » ou « 6100402 » pour « six cent quarante-deux »). Ainsi, avant l'apprentissage formel, les enfants ont acquis le sens de l'écriture des nombres : le premier chiffre mentionné s'écrit à gauche, le deuxième au milieu et le troisième à droite, etc. Ils ont également compris que le zéro par rapport aux autres chiffres possède un rôle particulier : représentant le néant mais marquant aussi une position.

Certains nombres écrits correctement semblent être acquis par imprégnation dû à leur fréquence dans la vie quotidienne (ex : 100). Il faut noter que la présence de l'écriture littérale est fréquente en GSM, mais qu'elle est signe de difficultés en mathématiques si elle persiste dans les classes de niveau supérieur, notamment en primaire (Moeller et al., 2011).

Les prémices observées sont confirmées par l'étude d'Herzog et al. (2019) qui a entrepris de mettre à jour les modèles de développement conceptuel de la VP existants (Cobb & Wheatley, 1988 ; Fuson, Wearne, et al., 1997 ; Ross, 1989 ; Van de Walle et al., 2004), en observant les compétences de 1300 enfants. Quatre phases sont décrites. Avant le premier stade, les enfants n'ont aucune conscience des valeurs des chiffres au sein du nombre. Les nombres leur apparaissent comme des entités unitaires. A ce niveau, ils sont capables de décomposer un nombre, mais n'utilisent pas particulièrement la base 10. Ce type de décomposition flexible, c'est-à-dire pouvoir dire que 14 c'est 8 et 6, mais aussi 12 et 2, etc., est d'ailleurs reconnu comme un principe essentiel à la VP (Ladel & Kortenkamp, 2016). Le prérequis de ce stade est l'acquisition du concept des chiffres de 1 à 9 (Cawley et al., 2007).

## **5.2. L'apprentissage formel**

Le premier stade du modèle d'Herzog et al. (2019) est atteint grâce à l'entrée dans l'apprentissage formel : les enfants commencent à distinguer les unités de regroupement par le vocabulaire « unité, dizaine, centaine ». Les enfants sont ainsi capables de nommer les positions des chiffres (ex : dire que 8 dans 82 est une dizaine). Cependant, pour identifier le chiffre des unités ou des dizaines, la compréhension de la relation entre les unités de regroupement n'est pas nécessaire. Ces unités de regroupement restent sans lien entre elles. A ce stade, les constructions et les décompositions de nombres sont canoniques, c'est-à-dire se basant sur des groupes de 10.

Au deuxième stade, les enfants comprennent la relation entre les dizaines et les unités s'ils disposent d'un support visuel. Le concept de la dizaine n'est pas assez élaboré pour former une unité de regroupement, ils ont besoin de vérifier la relation unité-dizaine en comptant les unités dans chaque dizaine. Les enfants peuvent manipuler des représentations non canoniques s'ils ont un support visuel pour structurer leurs regroupements. La compréhension de la VP semble essentiellement fonctionnelle à ce stade, c'est-à-dire suffisante pour la manipulation, mais non conceptuelle : non comprise.

Le troisième stade est atteint lorsque la relation unité-dizaine est suffisamment intériorisée pour pouvoir se passer de support visuel : les enfants n'ont plus besoin de vérifier l'équivalence de 10 unités dans une dizaine, elle n'est plus un simple symbole mais un ensemble de dix unités combinées en une unité abstraite. La manipulation de représentations non canoniques sans support visuel devient possible. Herzog et al. (2019) observent que l'abstraction de la relation entre dizaines et unités ne se généralise pas aux unités de regroupement supérieures comme la centaine ou le millier. Pour ce type de grandeurs, les enfants ont besoin à nouveau d'un support visuel pour construire et vérifier la relation entre les dizaines et les centaines par le comptage. Byrge et al. (2014) corroborent ces résultats en observant des enfants écrire « 70034 » ou « 710034 » pour « sept cent trente-quatre ».

Le quatrième et ultime stade est atteint lorsque la généralisation des regroupements est acquise aux grandeurs supérieures. Aucune limitation d'ordre de grandeur n'est observée. Le modèle s'achève à ce stade : la VP est acquise. Les principes fondamentaux de la compréhension approfondie de la VP sont la capacité de regroupement par 10, la connaissance des positions des chiffres dans le nombre et la capacité à composer et décomposer de manière flexible les nombres (Ladel & Kortenkamp, 2016 ; Van de Walle et al., 2004).

Ce modèle est théorique et doit être appuyé et confirmé par d'autres études. Certaines études sont concordantes, Chan et al. (2017) ont observé les compétences des enfants de CP sur une épreuve d'identification de quantités à partir de représentations par du matériel (petits cubes pour les unités, barres pour les dizaines et carrés pour les centaines). La tâche était d'écrire le nombre correspondant au matériel donné. Les auteurs ont repéré et analysé plusieurs erreurs correspondant au modèle : l'absence de valeur de position (ex : 5 dizaines et 4 unités comptées 9) correspondant au stade 0 ; des erreurs d'écriture terme à terme avec une valeur de position mais sans intégration des deux parties (ex : 504 pour 5 dizaines et 4 unités) correspondant au stade I ; des erreurs de regroupement lorsque les représentations ne sont pas canoniques (ex : 2 dizaines, 3 unités, 4 dizaines écrit 234), correspondant au stade II. D'autres erreurs ne s'intégrant pas au modèle ont été observées, comme des erreurs de comptage après le passage du 9 (ex : « 27, 28, 29, 40 »), qui soulignent la difficulté de la gestion du changement de dizaines. Le caractère commun de ces erreurs témoigne de la complexité de cet apprentissage. Le développement de

l'automatisation du traitement des nombres s'effectue plus tard dans les classes du CE1 au CM1 (Mann et al., 2012).

En résumé, dès la maternelle, les enfants ont une conception principalement unitaire du nombre et sont capables d'extraire intuitivement quelques régularités du système de numération indo-arabe. C'est par l'apprentissage formel que se développe la VP. La maîtrise de cette compétence est longue et semble s'étaler sur les classes de primaire. Ces éléments sont cependant trop minces pour constituer des repères de développement de cette acquisition. Cawley et al. (2007) soulignent que certains enfants n'acquièrent pas une compréhension approfondie de la VP et compensent par leurs capacités mnésiques. L'évaluation fine de la compréhension de VP est essentielle.

## **6 État des lieux des outils d'évaluant la valeur positionnelle**

### **6.1. Dans la recherche**

De nombreuses tâches expérimentales ont été mises au point dans le domaine de la recherche pour tester la maîtrise ou la compréhension de la VP. Premièrement, il existe des tâches de jugement. Il s'agit de comparaison de nombres en code indo-arabe avec modulation de la compatibilité entre dizaines et unités (par exemple 42 et 57 est une paire compatible car  $4 < 5$  et  $2 < 7$ , alors que 48 et 63 sont incompatibles ; Moeller et al., 2011), d'ordination de nombres indo-arabes (Mix et al., 2017), de décomposition de nombres (par exemple choisir si 27 est égal à  $2+7$ ,  $20+7$ ,  $20+70$ ,  $2+70$  ; Lafay et al., 2020). Il existe également une tâche de jugement d'opération avec retenue (Moeller et al., 2011). Ces tâches peuvent être qualifiées de mesures indirectes de la compréhension de la VP car elles mettent en lumière l'utilisation de la VP au sein d'une autre action. Une seule tâche de jugement parmi celles recensées est qualifiable de mesure directe, il s'agit de celle de jugement d'association d'une quantité analogique à un chiffre au sein d'un nombre (PicPVT ; Osana et al., 2018), la question portant directement sur la VP.

Deuxièmement, sont présentes les tâches demandant une production de l'enfant : dictée de nombres indo-arabes (Dietrich et al., 2016), création du plus grand ou plus petit nombre possible avec des chiffres donnés (Laski et al., 2016), comptage flexible (par exemple : de 10 en 10 ; Aunio & Räsänen, 2016).

Le troisième type de tâche est composé de manipulation d'objets concrets ou de représentations figurales de ces objets. Il peut s'agir de perles ou de cubes, dont le

dispositif complet est plus ou moins transparent et proportionnel. L'enfant doit représenter un nombre indo-arabe avec ces objets ou inversement, écrire le nombre correspondant aux objets donnés. La présentation des objets peut être canonique ou non (Chan, 2014 ; Chan et al., 2017 ; Lafay et al., 2020).

Enfin, sont présentes également des tâches de désignation : il s'agit d'identifier explicitement le chiffre des unités, dizaines ou centaines d'un nombre en code indo-arabe, ou d'identifier les objets représentant l'unité, la dizaine ou la centaine d'un nombre (Laski et al., 2016). Ces trois derniers types de tâches peuvent être qualifiés de mesures directes.

## **6.2. En clinique**

Trois batteries à disposition des orthophonistes proposent de tester la maîtrise de la compréhension du système de numération chez les enfants âgés de 5 à 8 ans (Lafay & Cattini, 2018).

La batterie Tedi-Math (Van Nieuwenhoven et al., 2001) propose un module nommé « Compréhension du système de numération » conçu en deux subtests. Le premier invite l'enfant à identifier le chiffre des unités puis celui des dizaines au sein de nombres de deux à trois chiffres. Dans le second, l'enfant doit, avec ou sans manipulation identifier un nombre de bâtonnets qui lui sont présentés groupés en paquet de dix ou seul. Puis, avec des pièces d'un ou dix euros, il doit représenter des sommes données. Il faut noter que ces épreuves ne sont présentées qu'à partir du CE1 et ne permettent donc pas une évaluation précoce de la compétence.

Au sein de son module de numération, la batterie B-LM (Métral, 2008) propose une tâche : l'enfant, après avoir dénombré une collection de jetons, doit écrire le nombre correspondant en code indo-arabe puis dire oralement ce qu'est une unité, une dizaine, montrer dans le nombre le chiffre des unités et celui des dizaines, et enfin donner le nombre de jetons correspondant au chiffre des unités et au chiffre des dizaines. Notons que cette épreuve comporte un seul item et la batterie ne fournit pas de normes.

La batterie présentée dans l'ouvrage *Difficultés en mathématiques, évaluation et rééducation* (Koppel, 1998) dispose d'une partie intitulée « Compréhension de la transcription du nombre au-delà de neuf ». Cette partie demande à l'enfant de montrer le chiffre des dizaines au sein d'un nombre ou, à l'inverse, de dire ce que représente un chiffre donné au sein du nombre présenté. Notons que cette épreuve comporte un seul item et la batterie ne fournit pas de normes.



De ces trois batteries, il est surtout question de tester le vocabulaire mathématique et non la compréhension approfondie de la VP. Il subsiste un écart trop important entre les outils déployés en recherche et ceux en clinique, aussi bien en termes de quantité que de qualité.

## **7 Propriétés psychométriques d'un outil d'évaluation**

Un outil d'évaluation doit répondre à certaines propriétés psychométriques pour être considéré de bonne qualité. Cet outil doit être standardisé, valide, fidèle et posséder des données normatives (Lafay & Cattini, 2018). La standardisation concerne l'uniformité des conditions de passation et de cotation afin de limiter l'effet de la subjectivité de l'administrateur du test.

### **7.1. Validité**

La validité fait référence au degré auquel un test mesure réellement ce qu'il prétend mesurer. La validité de surface (Ivanova & Hallowell, 2013) concerne la recevabilité du test par l'évaluateur et le patient, c'est-à-dire l'acceptation sociale du test aux yeux d'observateurs non-initiés. La validité de contenu (Bouchard et al., 2009) considère la pertinence du test en regard des modèles théoriques récents.

La validité de critère est la capacité à évaluer de manière appropriée la performance selon un critère de référence (Wainer & Braun, 1988). Elle est composée d'une part de la validité prédictive, c'est-à-dire une comparaison différée entre le test et un critère indicateur d'une performance future dans une tâche similaire que l'on cherche à prédire. D'autre part, elle concerne la validité concomitante (aussi appelée validité concourante par Leclercq et Veys, 2014). Cette dernière implique une comparaison au même moment de la mesure, entre l'outil et un critère de référence externe comme par exemple un autre test standardisé mesurant le même construit théorique.

La validité de construit se réfère à la capacité d'un test à mesurer un construit théorique. Elle est vérifiée par des analyses de la précision (sensibilité et spécificité), de la validité factorielle et de la validité en lien avec les caractéristiques des individus (Ivanova & Hallowell, 2013). Un outil est sensible lorsqu'il permet de classer une proportion importante de personnes en difficultés : il s'agit des vrais positifs (Plante & Vance, 1994). Un outil est spécifique lorsqu'il permet de classer une proportion importante de personnes ne présentant pas de difficultés : il s'agit des vrais négatifs. La validité factorielle est la mesure mettant en correspondance la structure théorique

du test et la structure statistique observée. Enfin, la validité en lien avec les caractéristiques des individus s'assure que, lorsque le construit mesuré est intrinsèquement lié à une ou plusieurs caractéristiques évidentes de l'individu, cette mesure de construit est sensible à cette relation. Il peut s'agir du genre, de l'âge, du niveau socio-économique, de la pathologie, ou encore du niveau scolaire. La recherche dans le cadre de ce mémoire s'intéressera aux validités concomitantes et de construit.

## **7.2. Fidélité**

La fidélité d'un test porte sur son degré de cohérence, de précision et de reproductibilité. Elle concerne la stabilité temporelle, inter-juges, et parallèle, ainsi que la cohérence interne, c'est-à-dire la cohérence du test avec lui-même : chaque composante du test réagit de manière cohérente à une même réponse. La cohérence interne est calculée par l'alpha de Cronbach (Nunnally, 1978), par un calcul de corrélation entre items ou inter-épreuves (Cohen, 1988) ou estimée par la corrélation d'une bissection des items. La recherche dans le cadre de ce mémoire s'intéressera à ce dernier type de fidélité : la cohérence interne.

## **7.3. Définition d'une norme critériée**

Une norme correspond à la distribution des scores obtenus par un échantillon de personnes, représentatif d'une population définie, à un instrument administré dans des conditions standardisées (Lafay & Cattini, 2018). Le qualificatif critérié renvoie à une notion de quantification et permet de mesurer une habileté particulière en termes de niveau de maîtrise (Glaser, 1963). Doit être défini un seuil de réussite ou de maîtrise par la moyenne et les écart-types des scores de l'échantillon.

## **8 Problématiques**

Trop peu d'outils évaluant la compréhension de la VP sont à la disposition des orthophonistes, et ceux-ci ne testent pas une maîtrise approfondie de la VP. Pourtant la compréhension de VP est essentielle à l'apprentissage des mathématiques, et est un précurseur des compétences arithmétiques futures. Il subsiste un réel besoin d'outils testant les capacités d'enfant en apprentissage. Ce nouvel outil, pour répondre à ce manque, doit être élaboré en s'inspirant des outils existants dans les études expérimentales mais doit être normé et étalonné à l'échelle d'une plus grande population pour être utilisables en clinique. Helloin et Lafay (en préparation) développent un module de numération au sein d'une batterie d'évaluation des

habiletés mathématiques de l'enfant entre 5 et 8 ans, correspondant aux classes de GSM jusqu'au CE1. Cette batterie est nommée Examath 5-8. Pour évaluer la compréhension du système de numération indo-arabe à valeur positionnelle, le module comporte des épreuves de représentation de quantités avec des objets virtuels à partir de nombres indo-arabes, et inversement l'écriture en code indo-arabe à partir des objets virtuels, ainsi qu'une tâche d'identification de la VP. Le module évalue aussi l'utilisation de la VP par des tâches plus indirectes de lecture et de dictée de nombres, et de décomposition de nombres. Enfin, les données recueillies sur ces nouvelles épreuves pourront permettre d'enrichir les données développementales de la recherche sur les capacités des enfants de ces niveaux scolaires.

## **9 Objectifs de recherche**

L'objectif général de ce mémoire est d'évaluer certaines composantes psychométriques du module de numération de la batterie d'évaluation Examath 5-8 et d'étudier les compétences des enfants entre 5 et 8 ans aux tâches évaluant la VP.

Le premier objectif spécifique de ce présent mémoire est d'évaluer la cohérence interne du module de numération de la batterie Examath 5-8. Selon l'hypothèse 1, les épreuves de la batterie d'Examath 5-8 seront corrélées entre elles.

Le deuxième objectif spécifique est d'évaluer la validité concomitante du module de numération de la batterie d'évaluation Examath 5-8. Selon l'hypothèse 2, les scores aux épreuves de transcodage d'Examath 5-8 seront corrélés aux scores des épreuves de transcodage de la batterie Tedi-Math (hypothèse 2a), les scores aux épreuves de compréhension du système de numération d'Examath 5-8 seront corrélés aux scores de l'épreuve Pic PVT et aux scores des épreuves sur le système de numération de la batterie Tedi-Math (hypothèse 2b).

Le troisième objectif spécifique est d'évaluer la validité de construit en lien avec les caractéristiques développementales des enfants du module de numération de la batterie d'évaluation Examath 5-8. Selon l'hypothèse 3, les enfants scolarisés en cours élémentaire de première année (CE1) obtiendront de meilleurs scores et temps de réponse que les enfants scolarisés en cours préparatoire (CP), qui eux-mêmes obtiendront de meilleurs scores et temps de réponses que les enfants scolarisés en grande section de maternelle (GSM), ce pour l'ensemble des tâches.

Enfin, le quatrième objectif spécifique est d'établir une norme critériée du développement de la compétence en numération : pour cela, nous analyserons si les

groupes d'enfants obtiennent un score suffisant pour considérer la notion acquise sur ces nouvelles épreuves, mais également les types d'erreurs commises par les enfants scolarisés dans chaque niveau.

## II. Méthode

### 1 Participants

Le recrutement a pris la forme d'une large prospection visant les écoles publiques et privées de Lyon. Deux écoles privées ont accueilli le projet au sein de leur établissement. Une notice d'information à destination des parents ainsi qu'un visuel ont été diffusés par les établissements aux parents d'élèves. De par le public accueilli dans ce type d'établissement, les participants appartenaient pour une grande majorité à une classe socio-professionnelle élevée. Les parents ou responsables légaux des enfants ayant participé à l'étude ont signé un document de non-opposition à la participation.

Les critères d'inclusion étaient : (a) être en classe de GSM, CP ou CE1, (b) ne pas bénéficier ou ne pas avoir bénéficié d'une prise en soin orthophonique concernant la cognition mathématique, (c) être scolarisé-e, (d) être francophone en première, seconde ou troisième langue.

Ont participé aux passations 15 enfants de GSM (âge :  $M = 5.4$  ;  $ET = 0.3$ ), 12 enfants de CP (âge :  $M = 6.8$  ;  $ET = 0.2$ ) et 14 enfants de CE1 (âge :  $M = 7.4$  ;  $ET = 0.3$ ).

### 2 Matériels

#### 2.1. Examath 5-8

Cinq épreuves issues de la batterie Examath 5-8 (Helloin & Lafay, en préparation) ont été utilisées.

##### 2.1.1 *Évaluation de la maîtrise des codes*

L'épreuve Transcodage en lecture de nombres a pour objectif l'évaluation du passage du code arabe au code oral (le stimulus est visuel, un nombre en écriture arabe apparaît à l'écran). Trente items sont proposés, allant progressivement de nombre d'un à quatre chiffres et organisés en trois sous-épreuves (nombres inférieurs à 30, nombres de 31 à 99, nombres à trois et quatre chiffres). La réponse est donnée oralement. Le critère d'arrêt est fixé à quatre erreurs consécutives. Un point est donné par bonne réponse. Le score de chaque sous-partie est sur 10, le total sur 30. Le temps moyen de réponse par item lu correctement est calculé.

L'épreuve Transcodage en dictée de nombres a pour objectif l'évaluation du passage du code oral au code arabe (le stimulus est auditif : une même voix humaine énonce chaque item). Les 30 items vont progressivement de nombre d'un à quatre chiffres, organisés en trois sous-épreuves (nombres inférieurs à 30, nombres de 31 à 99, nombres à trois et quatre chiffres). La réponse est écrite par l'enfant sur une feuille vierge. L'expérimentatrice saisit au clavier la réponse écrite par l'enfant. Un point est donné par bonne réponse. Le critère d'arrêt est fixé à quatre erreurs consécutives. Le score de chaque sous-partie est sur 10, le total sur 30.

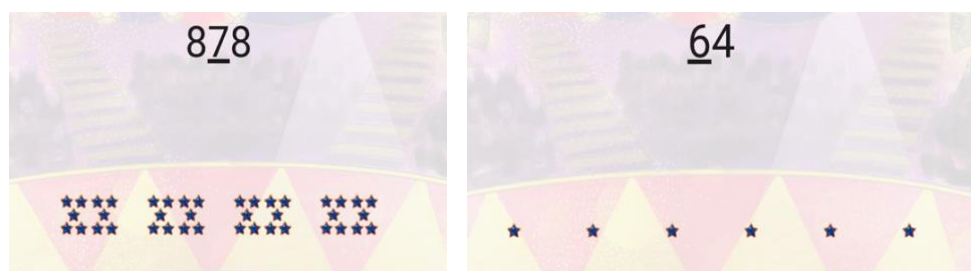
### 2.1.2 Évaluation de la compréhension du système de numération

Trois tâches ont été utilisées pour évaluer la compréhension et l'utilisation du système de numération indo-arabe.

L'épreuve Valeur des chiffres évalue l'acquisition de la connaissance et la compréhension de la valeur positionnelle et de la valeur multiplicative. Il s'agit de juger une association entre un chiffre souligné au sein d'un nombre et d'une quantité de points (code analogique). Deux exemples d'items sont proposés dans la Figure 1. L'enfant doit répondre oralement par « vrai » ou « faux ». Quinze items proposent des nombres à deux ou trois chiffres. Un point est donné par bonne réponse pour un score total sur 15.

**Figure 1**

*Capture d'Écran d'Items Proposés Dans le Test Valeur des Chiffres.*



*Légende : A gauche, l'erreur porte sur la valeur multiplicative. A droite, l'erreur porte sur la valeur positionnelle. Note : D'autres items ne comportent pas d'erreur.*

L'épreuve appelée Décomposition additive consiste à juger une décomposition additive de nombres présentés en code arabe (exemple :  $13 = 10 + 3$ ). L'enfant doit répondre oralement par « vrai » ou « faux ». Dix items sont proposés. Parmi les stimuli erronés, on retrouve des erreurs de juxtaposition (par exemple :  $54 = 5 + 4$ ), des erreurs liées aux irrégularités du code oral du français (par exemple :  $89 = 4 + 20 + 9$ ), des erreurs liées à une compréhension partielle de la VP (par exemple :  $182 = 10 +$

82). Enfin, certains items cumulent plusieurs types d'erreurs, comme l'item  $75 = 60 + 5$  qui compte une erreur de liée aux irrégularités du code oral et une erreur de calcul. Un point est donné par bonne réponse, permettant un score sur 10.

Enfin, une épreuve de production, nommée UDC, évalue la compréhension du système de numération à valeur positionnelle. Une première sous-épreuve appelée Arabe-Analogique demande à l'enfant de produire avec un matériel virtuel la représentation analogique du nombre présenté en code indo-arabe. Le matériel est représenté à l'écran soit par un cube pour l'unité et une barre pour la dizaine où apparaissent les dix cubes qui la composent, soit par une bille pour l'unité et une boîte de dix billes pour la dizaine. L'enfant montre à la testeuse le matériel à déposer dans une zone dédiée de l'écran. Six items de nombres à deux chiffres sont proposés. Un point est donné si le total est juste avec exclusivement les unités (par exemple : 13 unités pour la cible 13, et non une dizaine et trois unités), deux points sont accordés si le total est juste avec au moins une partie des dizaines (par exemple : deux dizaines et treize unités pour 33). Enfin, trois points sont accordés si le total est juste avec le nombre attendu de dizaines et d'unités. Le score total est sur 18. Dans la seconde sous-épreuve, appelée Analogique-Arabe, est présentée à l'enfant la représentation imagée du nombre dans le code analogique avec le même matériel virtuel que précédemment, et l'enfant doit écrire le nombre correspondant en code arabe sur un papier libre. Six items sont proposés, un point est donné pour chaque bonne réponse, le score total est sur 6. Le score total de l'épreuve UDC est sur 24.

## **2.2. Tedi-math**

Deux épreuves de la batterie Tedi-math (Van Nieuwenhoven et al., 2001) ont été utilisées pour évaluer la maîtrise des codes (transcodage en lecture de 20 nombres allant d'un à trois chiffres et transcodage en dictée de 20 nombres allant d'un à trois chiffres, scores respectifs sur 20 et 20). La compréhension du système de numération en base 10 est évaluée en quatre sous-épreuves. La première est une tâche de représentation de nombres avec du matériel concret (bâtonnets seuls ou attachés en groupe de 10), trois items sont proposés, pour 3 points. La seconde est une tâche de représentation mentale sans matériel. Par exemple, la testeuse demande à l'enfant : « Un paquet est toujours fait avec 10 bâtonnets. Si j'ai 14, je peux avoir combien de paquets et il restera combien de bâtonnets tout seuls ? ». Quatre items sont proposés, pour 4 points. La troisième sous-épreuve est une tâche de manipulation en représentation mentale. Par exemple la testeuse demande : « J'ai 15 bâtonnets. Je

veux donner 7 bâtonnets à mon ami. Est-ce que je dois ouvrir un paquet ou bien est-ce que j'ai assez de bâtonnets tout seuls ? Pourquoi ouvrir un paquet ? ». Quatre items sont proposés pour 4 points. Enfin, une tâche de connaissance du vocabulaire mathématique est proposée : l'enfant doit entourer le chiffre des unités (5 items, 5 points), puis le chiffre des dizaines (5 items, 5 points), sur des nombres écrits en code indo-arabe d'un à trois chiffres. Le score total du module « Compréhension du système de numération » est sur 21 points.

### **2.3. PicPVT**

La tâche Picture Place Value Task (PicPVT ; Osana et al., 2018) a pour objectif l'évaluation de la compréhension de la valeur positionnelle du système de numération. Elle consiste en un jugement d'associations entre un chiffre souligné au sein d'un nombre et une représentation analogique : groupe de points représentant une unité, une dizaine et une centaine. L'enfant répond oralement « vrai » ou « faux ». La tâche comporte 20 items qui pouvaient être juste ou portaient une erreur sur la valeur positionnelle (un point donné par bonne réponse, score total sur 20).

## **3 Procédure générale**

La passation des tests s'est déroulée au sein des établissements scolaires, sur le temps scolaire, dans une salle mise à disposition, en décembre 2020. Les passations étaient individuelles et menées par la même testeuse (l'autrice du présent mémoire). Les participant-e-s ont été vu-e-s une unique fois. Les règles sanitaires en vigueur liées à la covid-19 ont été respectées.

Les consignes de passation des batteries d'évaluation ont été scrupuleusement respectées. Toutes les données ont été recueillies anonymement.

L'ordre de passation des tâches a été randomisé pour prévenir les effets d'une épreuve à l'autre. La randomisation a été effectuée à l'aide du site Random.org. Ont été générées 45 suites (objectif initial du nombre de participants). Ces suites, composées des neuf épreuves, ont été construites avec les contraintes suivantes : elles devaient être uniques, ne pas faire suivre les deux épreuves de dictée, les deux épreuves de lecture et enfin les deux épreuves de jugement d'association en code analogique et arabe.

### III. Résultats

Les analyses statistiques ont été réalisées avec le logiciel SPSS version 22.

#### 1 Hypothèse 1 – Cohérence interne

Les scores de chaque épreuve d'Examath 5-8 sont corrélés entre eux. La lecture et la dictée ( $r^2 = .98, p < .001$ ), la lecture et la tâche Décomposition d'addition ( $r^2 = .80, p < .001$ ), la lecture et la tâche Valeur des chiffres ( $r^2 = .75, p < .001$ ), la lecture et la tâche UDC ( $r^2 = .88, p < .001$ ), la dictée et la tâche UDC ( $r^2 = .88, p < .001$ ), la dictée et la tâche Valeur des chiffres ( $r^2 = .74, p < .001$ ), la dictée et la tâche Décomposition d'addition ( $r^2 = .79, p < .001$ ), la tâche UDC et la tâche Valeur des chiffres ( $r^2 = .68, p < .001$ ), la tâche UDC et la tâche Décomposition additive ( $r^2 = .73, p < .001$ ), et enfin la tâche Valeur des chiffres et la tâche Décomposition additive ( $r^2 = .68, p < .001$ ). Les effets sont de grandes tailles, les corrélations sont fortes. Les temps des épreuves de lecture et de dictée sont corrélés significativement ( $r^2 = .46, p = .003$ ), la corrélation est modérée.

#### 2 Hypothèse 2 – Validité concomitante

##### 2.1. Épreuve de Transcodage en Lecture

Le score total de la tâche en lecture d'Examath 5-8 est corrélé significativement au score de l'épreuve de lecture de Tedi-Math ( $r^2 = .97, p < .001$ ). Chaque sous-catégorie de score (items de 1 à 30, items de 31 à 99, items supérieurs à 99) est corrélée significativement à l'épreuve de lecture de Tedi-math (respectivement  $r^2 = .81, p < .001$  ;  $r^2 = .93, p < .001$  ;  $r^2 = .89, p < .001$ ). Les effets sont de grandes tailles, les corrélations sont fortes.

##### 2.2. Épreuve de Transcodage en Dictée

Le score total de la tâche en dictée d'Examath 5-8 est corrélé significativement au score de l'épreuve de dictée de Tedi-Math ( $r^2 = .98, p < .001$ ). Chaque sous-catégorie de score (items de 1 à 30, items de 31 à 99, items supérieurs à 99) est corrélée significativement à l'épreuve de dictée Tedi-math (respectivement  $r^2 = .83$  ;  $r^2 = .94$  ;  $r^2 = .86$  ;  $p < .001$  pour chaque). Les effets sont de grandes tailles, les corrélations sont fortes.

##### 2.3. Épreuve Valeur des chiffres

Le score de la tâche Valeur des chiffres d'Examath 5-8 est corrélé significativement au score de la tâche PicPVT ( $r^2 = .94, p < .001$ ), ainsi qu'à la tâche



Compréhension du système de numération de Tedi-Math ( $r^2 = .78, p < .001$ ). Les effets sont de grandes tailles, les corrélations sont fortes.

## 2.4. Épreuve UDC

Le score total de la tâche UDC d'Examath 5-8 est corrélé significativement au score de la tâche Compréhension du système de numération de la batterie Tedi-Math ( $r^2 = .86, p < .001$ ). Les sous-catégories de score Arabe-Analogique et Analogique-Arabe sont corrélées significativement à l'épreuve Tedi-math (respectivement  $r^2 = .85 ; r^2 = .89 ; p < .001$  pour chaque). Les effets sont de grandes tailles, les corrélations sont fortes.

## 3 Hypothèse 3 – Validité de construit en lien avec les caractéristiques des individus : le niveau scolaire

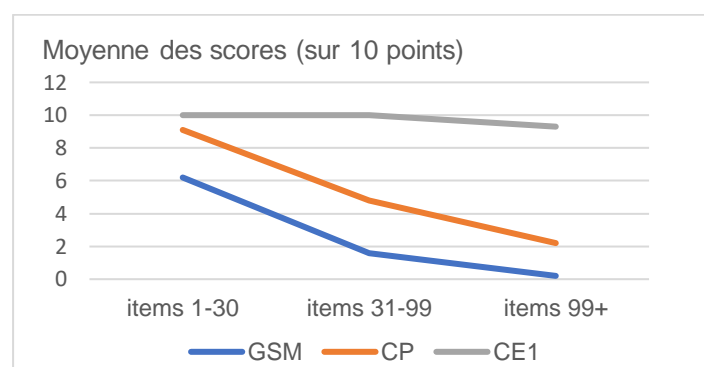
Une analyse de variance a été réalisée sur les variables dépendantes (score et temps à chaque épreuve) selon la variable indépendante Classe (GSM, CP, CE1).

### 3.1. Épreuve de Transcodage en Lecture

En lecture, les analyses montrent un effet significatif de la classe sur le score total ( $F(2,38) = 52.43, p < .001, \eta^2_p = .73$ ). Les analyses de comparaison montrent que les CE1 obtiennent un meilleur score que les CP et les GSM ( $p < .001$  et  $p < .001$ ), que les CP obtiennent un meilleur score que les GSM ( $p = .002$ ). Une analyse similaire a été réalisée pour les scores aux trois sous-épreuves et montre un effet significatif de la classe sur chacun ( $F(2,38) = 15.53, p < .001, \eta^2_p = .45 ; F(2,38) = 30.18, p < .001, \eta^2_p = .61 ; F(2,38) = 67.99, p < .001, \eta^2_p = .78$ ). La Figure 2 se propose d'illustrer les résultats.

### Figure 2

*Moyennes des Scores en Lecture en Fonction des Types d'Items, Selon les Classes.*



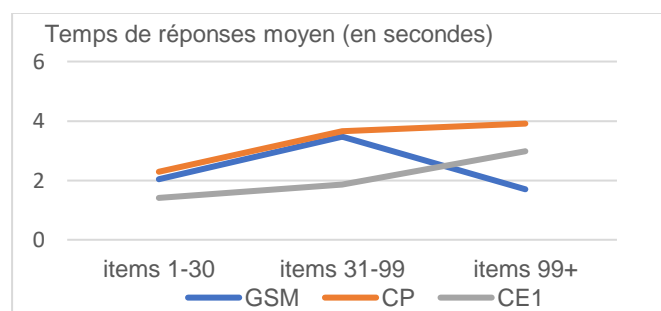
Pour les scores des items compris en 1 et 30, les résultats ne montrent pas de différence significative entre les CE1 et les CP ( $p = .69$ ), mais une différence

significative entre les CP et les GSM et entre les CE1 et les GSM (respectivement  $p = .001$ ,  $p < .001$ ). Pour les scores des items compris entre 31 et 99, toutes les classes obtiennent des scores significativement différents entre elles (entre les CE1 et les CP :  $p < .001$  ; entre les CE1 et les GSM :  $p < .001$  ; et enfin entre les CP et les GSM :  $p = .02$ ). Pour les scores des items supérieurs 99, les analyses montrent une différence significative entre les CE1 et les CP et entre les CE1 et les GSM (respectivement  $p < .001$  et  $p < .001$ ), mais pas de différence significative entre les CP et les GSM ( $p = .08$ ).

Concernant le temps, les analyses montrent un effet significatif de la classe sur le temps total ( $F(2,38) = 6.21$ ,  $p = .005$ ,  $\eta^2_p = .25$ ). Les analyses de comparaison montrent que le temps de réponse des CE1 est plus court que celui des CP ( $p = .007$ ) mais pas plus court que celui des GSM ( $p = 1.00$ ). Les CP ont un temps de réponse plus long que les GSM ( $p = .002$ ). Une analyse similaire a été réalisée pour les scores aux trois sous-épreuves et montre un effet significatif de la classe sur chacun ( $F(2,20) = 8.47$ ,  $p = .002$ ,  $\eta^2_p = .46$  ;  $F(2,20) = 17.77$ ,  $p < .001$ ,  $\eta^2_p = .64$  ;  $F(2,20) = 3.57$ ,  $p = .05$ ,  $\eta^2_p = .26$ ). Pour les temps des items compris en 1 et 30, les résultats montrent une différence significative entre les CE1 et les CP et entre les CE1 et les GSM (respectivement  $p = .01$  et  $p = .02$ ), mais une différence non significative entre les CP et les GSM ( $p = 1.00$ ). Pour les temps des items compris entre 31 et 99, les résultats montrent également une différence significative entre les CE1 et les CP et entre les CE1 et les GSM (respectivement  $p < .001$  et  $p = .02$ ), mais une différence non significative entre les CP et les GSM ( $p = .66$ ). Pour les temps des items supérieurs 99, les analyses ne montrent pas de différence significative entre les CE1 et les CP et entre les CE1 et les GSM (respectivement  $p = .36$  et  $p = .32$ ), mais une différence significative entre les CP et les GSM ( $p = .05$ ). La Figure 3 illustre ces résultats.

### Figure 3

*Moyennes des Temps (en Seconde) à l'Épreuve de Lecture par Catégories d'items et Selon la Classe*



### 3.2. Épreuve de Transcodage en Dictée

Les analyses montrent un effet significatif de la classe sur le score total ( $F(2,38) = 49.16, p < .001, \eta^2_p = .72$ ). Les analyses de comparaison montrent que les CE1 obtiennent un meilleur score que les CP et les GSM ( $p < .001$  et  $p < .001$ ) et que les CP obtiennent un meilleur score que les GSM ( $p = .01$ ). Une analyse similaire a été réalisée pour les scores aux trois sous-épreuves et montre un effet significatif de la classe sur chacun ( $F(2,38) = 19.09, p < .001, \eta^2_p = .50$  ;  $F(2,38) = 22.86, p < .001, \eta^2_p = .55$  ;  $F(2,38) = 81.11, p < .001, \eta^2_p = .81$ ). Pour les scores des items compris en 1 et 30, les résultats ne montrent pas de différence significative entre les CE1 et les CP ( $p = .27$ ), mais une différence significative entre les CE1 et les GSM et entre les CP et les GSM (respectivement  $p < .001, p = .001$ ). Pour les scores des items compris entre 31 et 99, les analyses montrent une différence significative entre les CE1 et les CP et entre les CE1 et les GSM (respectivement  $p = .001$  et  $p < .001$ ), mais pas de différence significative entre les CP et les GSM ( $p = .07$ ). Pour les scores des items supérieurs 99, les analyses montrent une différence significative entre les CE1 et les CP et entre les CE1 et les GSM (respectivement  $p < .001$  et  $p < .001$ ), mais pas de différence significative entre les CP et les GSM ( $p = .17$ ).

Concernant le temps, les analyses montrent un effet significatif de la classe sur le temps total ( $F(2,38) = 3.88, p = .03, \eta^2_p = .17$ ). Les analyses de comparaison montrent que le temps de réponse des CE1 n'est pas significativement plus court que celui des CP ( $p = .14$ ) mais plus court que celui des GSM ( $p = 0,04$ ). Le temps de réponse des CP n'est pas significativement plus court que celui de la classe des GSM ( $p = 1.00$ ). Une analyse similaire a été réalisée pour les scores aux trois sous-épreuves et montre un effet significatif de la classe pour les temps des items compris entre 1 et 30 et pour les items compris entre 31 et 99 (respectivement  $F(2,23) = 29.44, p < .001, \eta^2_p = .72$  ;  $F(2,23) = 33.16, p < .001, \eta^2_p = .74$ ), mais pas pour les items supérieurs à 99 ( $F(2,23) = 1.04, p = .37, \eta^2_p = .08$ ). Pour les temps des items compris en 1 et 30, toutes les classes ont des temps significativement différents entre elles (entre les CE1 et les CP :  $p = .01$  ; entre les CE1 et les GSM :  $p < .001$  ; et enfin entre les CP et les GSM :  $p = .001$ ). Pour les temps des items compris entre 31 et 99, toutes les classes ont également des temps significativement différents entre elles (entre les CE1 et les CP :  $p < .001$  ; entre les CE1 et les GSM :  $p < .001$  ; et enfin entre les CP et les GSM :  $p = .02$ ).

### 3.3. Épreuve UDC

Les analyses montrent un effet significatif de la classe sur le score ( $F(2,38) = 30.29, p < .001, \eta^2_p = .61$ ). Les analyses de comparaison montrent que toutes les classes obtiennent un score significativement différent (entre les CE1 et les CP ( $p = .002$ ), entre les CE1 et les GSM ( $p < .001$ ), entre les CP et les GSM ( $p = .03$ ). Une analyse similaire a été réalisée pour les scores aux deux sous-épreuves et montre un effet significatif de la classe sur chacun ( $F(2,38) = 26.72, p < .001, \eta^2_p = .58$  ;  $F(2,38) = 35.68, p < .001, \eta^2_p = .65$ ). Les analyses de comparaison montrent, pour les items de la sous-épreuve « Arabe-analogie », une différence significative entre toutes les classes : entre les CE1 et les CP ( $p = .002$ ), entre les CE1 et les GSM ( $p < .001$ ) et entre les CP et les GSM ( $p = .01$ ). Pour les items de la sous-épreuve « analogie-arabe », toutes les classes obtiennent des scores significativement différents : entre les CE1 et les CP ( $p = .01$ ), entre les CE1 et les GSM ( $p < .001$ ) et entre les CP et les GSM ( $p < .001$ ).

### 3.4. Épreuve Valeur des chiffres

Les analyses montrent un effet significatif de la classe sur le score ( $F(2,38) = 46.19, p < .001, \eta^2_p = .71$ ). Les analyses de comparaison montrent des scores significativement différents entre les CE1 et les CP ainsi qu'entre les CE1 et les GSM (respectivement  $p < .001$  et  $p < .001$ ). Cependant les scores des CP et GSM ne sont pas significativement différents ( $p = .31$ ).

### 3.5. Épreuve Décomposition d'addition

Les analyses montrent un effet significatif de la classe sur le score ( $F(2,38) = 40.77, p < .001, \eta^2_p = .68$ ). Les analyses de comparaison montrent que toutes les classes obtiennent un score significativement différent (entre les CE1 et les CP ( $p < .001$ ), entre les CE1 et les GSM ( $p < .001$ ) et entre les CP et les GSM,  $p = .05$ ).

L'ensemble des résultats est illustré en annexe A.

## 4 Hypothèse 4 – Norme critériée

Les performances des participants (moyennes et écart-types des scores) aux épreuves de la batterie d'Examath 5-8 ont permis d'élaborer une norme critériée des compétences pour chaque niveau scolaire et pour chaque épreuve. Le critère choisi pour considérer une épreuve réussie et inférer une notion acquise est le seuil de 75%. Le critère choisi pour inférer une notion en cours d'acquisition est le seuil de 60 %. Les résultats aux épreuves de Lecture et Dictée de nombres sont recensés dans le

Tableau 1. Les nombres de 1 à 30 sont à considérer en cours d'acquisition pour les GSM et acquis pour les CP et les CE1. Les nombres supérieurs à 31 ne sont acquis qu'en classe de CE1. Il n'y a pas de variabilité entre les deux modalités lecture et dictée.

**Tableau 1**

*Norme Critériée des Compétences en Lecture et Dictée Attendues Selon la Classe*

		LECTURE			DICTÉE		
		Score_1_30	Score_31_99	Score_99plus	Score_1_30	Score_31_99	Score_99plus
<b>GSM</b>	Moyenne	6,2 /10	1,6 /10	0,2 /10	6,5 /10	2,3 /10	0,3 /10
	Écart-type	2,8	3	0,4	2,1	3,4	0,5
		<b>En cours</b>	<b>Non Acquis</b>	<b>Non Acquis</b>	<b>En cours</b>	<b>Non Acquis</b>	<b>Non Acquis</b>
<b>CP</b>	Moyenne	9,1 /10	4,8 /10	2,2 /10	8,9 /10	5,1 /10	1,8 /10
	Écart-type	1,6	4,3	3,6	1,7	4,2	3,2
		<b>Acquis</b>	<b>Non Acquis</b>	<b>Non Acquis</b>	<b>Acquis</b>	<b>Non Acquis</b>	<b>Non Acquis</b>
<b>CE1</b>	Moyenne	10 /10	10 /10	9,3 /10	10 /10	9,9 /10	9,1 /10
	Écart-type	0	0	1,6	0	0,3	1,6
		<b>Acquis</b>	<b>Acquis</b>	<b>Acquis</b>	<b>Acquis</b>	<b>Acquis</b>	<b>Acquis</b>

Les résultats du Tableau 2 indiquent que les épreuves UDC et Décomposition d'addition ne sont acquises qu'en CE1. En revanche l'épreuve Valeur des chiffres est acquise dès le CP et elle est en cours d'acquisition chez les GSM. Une analyse détaillée des scores de la première partie de l'épreuve UDC a permis de montrer que 19 % des GSM réussissaient l'épreuve, contre 55 % des CP et 99 % des CE1.

**Tableau 2**

*Norme Critériée des Épreuves du Module de Numération d'Examath 5-8 Selon la Classe*

		UDC	Valeur des chiffres	Décomposition d'addition
<b>GSM</b>	Moyenne	4,4/24	10,8/15	4,5/10
	Écart-type	7,6	1,4	1,1
		<b>Non Acquis</b>	<b>En cours</b>	<b>Non Acquis</b>
<b>CP</b>	Moyenne	13,7/24	11,5/15	5,9/10
	Écart-type	8,9	0,7	2,2
		<b>Non Acquis</b>	<b>Acquis</b>	<b>Non Acquis</b>
<b>CE1</b>	Moyenne	23,7/24	14,5/15	9,3/10
	Écart-type	0,8	0,9	0,9
		<b>Acquis</b>	<b>Acquis</b>	<b>Acquis</b>

## 5 Analyse qualitative des erreurs

Une analyse des types de réponses et des erreurs produites par les enfants a été réalisée pour chaque épreuve, excepté pour l'épreuve Valeur des chiffres qui, par sa composition, ne permet pas d'interpréter la nature de l'erreur.

## 5.1. Erreurs de transcodage

Les erreurs ont été codées selon qu'elles étaient lexicales (erreur de position, inversion unité/dizaine) ou syntaxiques (ajout ou suppression de chiffre). Les analyses mettent en évidence une majorité nette d'erreurs lexicales en lecture et en dictée quelle que soit la classe (Tableau 3). Une exception est notable : les CE1 produisent autant d'erreurs lexicales que syntaxiques en dictée ; cela pourrait s'expliquer par la réalisation de la troisième partie de l'épreuve contenant des nombres à trois et quatre chiffres et « offrant » plus d'opportunités de produire des erreurs syntaxiques.

**Tableau 3**

*Pourcentage d'Erreurs de Type Lexicales ou Syntaxiques en Lecture et Dictée de Nombre Selon les Classes*

	LECTURE		DICTÉE	
	Lexicale	Syntaxique	Lexicale	Syntaxique
<b>GSM</b>	84 %	16 %	73 %	27 %
<b>CP</b>	100 %	0 %	83 %	17 %
<b>CE1</b>	100 %	0 %	50 %	50 %

## 5.2. Erreurs de valeur positionnelle

L'analyse de l'acceptation ou du rejet des items erronés de l'épreuve Décomposition d'addition permet d'observer que les erreurs liées à l'irrégularité du code oral ( $4 + 20 + 9 = 89$  ;  $60 + 5 = 75$  ;  $8 + 10 = 90$ ) sont rejetées massivement (donc bien repérées) par toutes les classes. Les erreurs concernant la VP ( $8 + 21 = 821$  ;  $5 + 4 = 54$ ) sont largement acceptées (et donc non repérées) par les GSM (80 %) et les CP (92 %). L'item «  $10 + 82 = 182$  » qui présente une erreur de VP avec le rôle particulier du zéro ne départage pas les comportements des classes de GSM et CP. Les résultats sont présentés dans le Tableau 4.

**Tableau 4**

*Pourcentage de Participants Rejetant ou Acceptant l'Item Erroné par Classe à l'Épreuve Décomposition d'Additions*

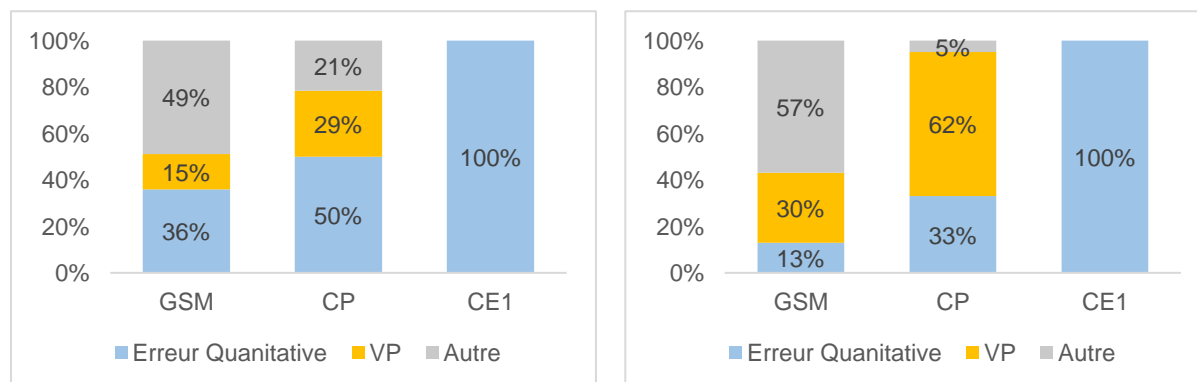
	<b>5 + 4 = 54</b>		<b>60 + 5 = 75</b>		<b>4 + 20 + 9 = 89</b>		<b>8 + 10 = 90</b>		<b>10 + 82 = 182</b>		<b>8 + 21 = 821</b>	
	Accepte Erreur	Rejet Erreur	Accepte Erreur	Rejet Erreur	Accepte Erreur	Rejet Erreur	Accepte Erreur	Rejet Erreur	Accepte Erreur	Rejet Erreur	Accepte Erreur	Rejet Erreur
<b>GSM</b>	80 %	20 %	20 %	80 %	20 %	80 %	20 %	80 %	33 %	67 %	80 %	20 %
<b>CP</b>	92 %	8 %	8 %	92 %	17 %	83 %	0 %	100 %	42 %	58 %	92 %	8 %
<b>CE1</b>	7 %	93 %	0 %	100 %	0 %	100 %	21 %	79 %	21 %	79 %	21 %	79 %

*Légende : en vert, la proportion majoritaire rejetant l'erreur ; en rouge, la proportion majoritaire acceptant l'erreur.*

Les erreurs commises lors de l'épreuve UDC pour les deux parties sont codées comme suit : « erreurs quantitatives » lorsqu'elles portent uniquement sur la dizaine ou l'unité (exemple : 2 dizaines et 5 unités pour la cible 24, ou 4 dizaines et 2 unités pour la cible 32) ; « VP » lorsque les catégories unité/dizaine ne sont pas considérées (exemple : 6 unités pour 33), lors d'inversion des dizaines et unités (exemple : 2 dizaines et 3 unités pour 32) ou encore lorsque la VP est partiellement respectée (exemple : 2 dizaines et 13 unités pour 33). Enfin, les erreurs ne pouvant être interprétées sont codées « Autre » (résultats présentés en Figure 4). Les GSM commettent une majorité d'erreurs Autres (49 % à la première sous-partie, 57 % à la deuxième). Ils font plus d'erreurs de VP dans le sens analogique-arabe (30 %) que dans le sens arabe-analogique (15 %). Les CP font plus d'erreurs de VP que les GSM et considérablement moins d'erreurs quantitatives (21 % sous-partie 1, 5 % sous-partie 2). A l'instar des GSM, les CP font plus d'erreurs de VP dans le sens analogique-arabe (62 %) que dans le sens arabe-analogique (29 %). Enfin, les CE1 ne commettent aucune erreur de VP, uniquement des erreurs quantitatives.

**Figure 4**

*Proportion du Type d'Erreur à l'Épreuve UDC en Fonction de la Classe*



*Légende : à gauche la première partie de l'épreuve UDC : représentation arabe vers la représentation analogique ; à droite la seconde partie de l'épreuve : représentation analogique vers la représentation arabe.*

## IV. Discussion

### 1 Interprétations des résultats

#### 1.1. Hypothèse 1 – Cohérence interne du module de numération

L'hypothèse 1 est validée : les épreuves du module de numération d'Examath 5-8 sont corrélées entre elles, le module possède une bonne cohérence interne. Les épreuves sont homogènes : les épreuves testent des construits différents mais qui possèdent des construits reliés et les compétences des enfants testés en numération se reflètent uniformément parmi les épreuves du module.

#### 1.2. Hypothèse 2 – Validité concomitante du module de numération

L'hypothèse 2 est validée : les épreuves permettant de vérifier la maîtrise du code, le transcodage en lecture et en dictée, sont corrélées aux épreuves similaires de la batterie Tedi-Math. L'épreuve Valeur des chiffres est corrélée à l'épreuve PicPVT. L'épreuve UDC est corrélée à l'épreuve Compréhension du système de numération de Tedi-Math. Le module de numération de la batterie d'Examath 5-8 possède une bonne validité concomitante : les résultats obtenus par cette batterie sont similaires à ceux obtenus par une batterie déjà existante. Un enfant obtenant un bon score avec la batterie Tedi-Math aura aussi un bon score avec Examath 5-8.

#### 1.3. Hypothèse 3 – Validité de construit

L'hypothèse 3 est validée. L'épreuve de Lecture de nombres permet de mettre en évidence une évolution des performances entre la GSM et le CE1 : les CE1 sont meilleurs que les CP qui eux-mêmes sont meilleurs que les GSM. Dans le détail, les CE1 et les CP ont des compétences similaires sur les nombres de 1 à 30 qui posent encore difficulté aux GSM. Au contraire, les CE1 sont meilleurs pour lire les nombres supérieurs à 99 que les CP et les GSM qui sont plus en difficulté pour ces nombres. Les nombres entre 31 et 99 sont les plus discriminants car toutes les classes obtiennent des scores significativement différents. La mesure des temps totaux de lecture permet aussi de mettre en évidence une évolution de la vitesse de lecture des nombres : chaque classe augmente son temps de réponse selon la grandeur du nombre à lire, sauf les GSM qui ont un temps plus court pour les nombres supérieurs à 99. En regard des scores faibles, ceci s'explique par le fait qu'ils ne sont pas familiers des nombres à trois chiffres et ne parviennent pas à les lire au complet. Chez les CP, l'augmentation du temps de réponse est nette dès les nombres de 31 à 99 alors que chez les CE1, elle a lieu à partir des nombres supérieurs à 99. Ces résultats



témoignent du degré d'automatisation de lecture des nombres selon le niveau scolaire : les CP ont automatisé la lecture des nombres de 1 à 30, les CE1 ont automatisé la lecture des nombres à deux chiffres jusqu'à 99.

Une évolution similaire par classe est également visible en Dictée : sur l'ensemble de l'épreuve, les CE1 sont meilleurs que les CP qui eux-mêmes sont meilleurs que les GSM. La compétence en CP se rapproche de celle en CE1 pour les nombres compris entre 1 et 30. En revanche, sur les nombres entre 31 et 99, là où toutes les classes se différencient en lecture, les CP et les GSM montrent des compétences similaires en dictée. Cette tendance est également visible pour les nombres supérieurs à 99. En CP, la lecture est donc maîtrisée avant le transcodage écrit. Les temps montrent la même évolution : les CE1 sont meilleurs que les GSM, mais les CP ne se démarquent pas significativement des CE1 et des GSM. Pour les nombres compris entre 1 et 30, les CE1 et les CP sont aussi rapides, les GSM plus lents. En revanche, pour les nombres compris entre 31 et 99, les CP et les GSM sont similairement plus lents que les CE1. Il n'y a pas d'effet de classe pour les nombres supérieurs à 99, les résultats ne sont pas interprétables.

Ces résultats montrent, d'une part, que la maîtrise du système de numération est un apprentissage long s'étalant sur deux années au minimum et, d'autre part, que les épreuves de Lecture et Dictée de nombres d'Examath 5-8 permettent de mettre en évidence l'évolution des performances de lecture et dictée de nombres des enfants en début d'apprentissage formel.

Concernant les épreuves évaluant la compréhension de la VP, pour l'épreuve UDC, l'hypothèse est validée : l'épreuve est discriminante. Toutes les classes obtiennent des scores différents selon la classe. Les CE1 sont meilleurs que les CP qui sont eux-mêmes meilleurs que les GSM, peu importe le sens de la tâche (du code analogique vers l'indo-arabe ou du code indo-arabe vers l'analogique). Pour l'épreuve Décomposition d'additions, l'hypothèse est aussi confirmée : les CE1 sont meilleurs que les CP qui sont eux-mêmes meilleurs que les GSM. Enfin, pour l'épreuve Valeur des chiffres, l'hypothèse est confirmée également : il y a un effet de la classe, les CE1 sont meilleurs que les CP et les GSM mais les CP et les GSM obtiennent des scores similaires. En regard de la norme critériée et de l'analyse des erreurs, plusieurs pistes d'interprétations seront proposées.

En résumé, ces résultats montrent que l'acquisition de la VP a lieu sur la période scolaire étudiée : entre la GSM et le CE1. Les CP obtiennent toujours un meilleur score que les GSM, marquant l'effet de l'apprentissage explicite.

#### **1.4. Hypothèse 4 – Établissement d'une norme critériée et analyse qualitative des erreurs**

Les productions des enfants aux tâches évaluant la VP, et surtout leurs erreurs, ont permis d'observer les compétences à chaque niveau et leur cohérence par rapport aux modèles de développement de compréhension de la VP mis au point par les chercheurs.

La connaissance de la VP des chiffres au sein des nombres est mesurée par l'épreuve Valeur des chiffres. Les CP et les GSM obtiennent des scores non significativement différents. Cela suggère plusieurs possibilités d'interprétation : (1) l'apprentissage qu'ont eu les CP ne leur permet pas de se différencier suffisamment des GSM sur cette épreuve, ou (2) l'épreuve n'est pas assez discriminante : les questions en réponses fermées (Vrai/Faux) offrent une possibilité de hasard favorable aux GSM pour rattraper les CP. En regard de la norme critériée établie : les GSM sont en cours d'acquisition sur cette tâche (60 à 75 % de réussite), alors que les CP et les CE1 l'ont acquise (plus de 75 % de réussite). Ceci suggère que cette tâche, qui mesure la connaissance de la VP sans avoir recours au vocabulaire, peut être réussie avant l'acquisition complète et approfondie du concept de la VP. Cette tâche mesurerait les prémices du développement de la compétence, avant l'apprentissage formel du vocabulaire mathématique.

Les productions des enfants sont observables par la tâche UDC. La norme critériée met en lumière qu'une petite proportion des GSM a réussi cette épreuve, contre la moitié des CP et la grande majorité des CE1. Ces proportions soutiennent le modèle d'Herzog et al. (2019) : les enfants en début d'apprentissage peuvent réussir l'épreuve avec un modèle visuel leur permettant de vérifier et soutenir la relation « dix unités correspond à une dizaine ». Ainsi, l'aspect procédural de la VP serait acquis avant l'acquisition conceptuelle : il s'agit pour la petite proportion de GSM et la moitié des CP observés d'une utilisation uniquement fonctionnelle de la VP avec support, sans compréhension approfondie du concept. L'analyse des erreurs montre une évolution des compétences procédurales de la VP : l'apprentissage explicite de la VP en CP semble réduire la proportion des erreurs non interprétables (codées « Autres ») et donc stabiliser dès la première année les productions, en comparaison avec les

GSM. Chez les GSM et les CP, on remarque une proportion moins importante d'erreurs quantitatives, c'est-à-dire d'erreurs de calculs ou d'estimation, dans le sens de traduction du code analogique vers le code indo-arabe. Il se pourrait que la représentation analogique facilite les manipulations de grandeurs importantes par la visualisation, soutenant les compétences d'estimation, de comptage et de dénombrement. Ainsi, en CP, la diminution des erreurs quantitatives permet une augmentation très importante de la proportion des erreurs de VP. Une interprétation probable est que les stratégies de regroupement à 10, non stables à ce stade, soient la principale source d'erreurs des CP : ceux-ci ne peuvent pas compter les dizaines directement. Ces dernières doivent faire l'objet d'une vérification de leur composition systématiquement, augmentant le risque d'erreur. Cependant, en GSM, les erreurs quantitatives diminuant dans le sens code analogique vers le code indo-arabe, c'est surtout la proportion des erreurs Autres qui augmente. Ceci permet d'inférer que les GSM, pour gérer et structurer des quantités analogiques, opèrent très peu de regroupement à 10, comme le suggère le modèle d'Herzog et al. (2019). Le regroupement arbitraire des dix cubes en une barre, présent dans l'épreuve, pourrait ne pas être suffisamment transparent pour être perçu – et utilisé – par les enfants de GSM. Ne pas utiliser la stratégie des regroupements à 10 pour les grandes quantités présentes dans l'épreuve engendre plus d'erreurs. Leurs productions erronées sont très éloignées de la cible, plus difficilement interprétables, et donc codées Autres.

La sensibilité des enfants aux règles qui régissent le système de numération est observable par la tâche Décomposition d'addition. Les CP obtiennent un score significativement différent et meilleur que les GSM, témoignant que les CP - en cours d'acquisition de la VP grâce à l'apprentissage formel - sont plus sensibles que les GSM à la construction du nombre. Ces résultats confirment que la compréhension de la VP est nécessaire à la construction du nombre et de sa décomposition. La norme critériée indique que seuls les CE1 ont acquis la tâche, les CP et les GSM ne l'ont pas acquise. Ces résultats sont appuyés par l'analyse des erreurs : les CP et les GSM sont très sensibles aux erreurs de valeurs de position (ex :  $5 + 4 = 54$ ), témoignant de la non-acquisition du concept de la VP. Une hypothèse expliquant ce phénomène serait que les enfants sont sensibles à l'ordre de grandeur des nombres. Dans l'item erroné «  $5 + 4 = 54$  », il y a autant de chiffres de part et d'autre de l'égalité, alors que pour l'item «  $10 + 82 = 182$  », il s'agit également d'une erreur de VP, mais avec un déséquilibre du nombre de chiffres de part et d'autre de l'égalité. En revanche, aucune des classes

n'est sensible aux erreurs liées aux irrégularités du code linguistique en français. La complexité linguistique du français concernant les mots-nombres parfois en partie non transparents ne semble pas constituer une difficulté majeure dans cette acquisition pour les enfants au développement typique. Il est certain que la connaissance du système de numération des CE1, grâce à l'acquisition de la VP, réduit leur sensibilité à la construction des mots-nombres irréguliers. Cependant, concernant les GSM, qui n'ont pas acquis la VP et qui ne sont pas familiers des grands nombres à deux ou trois chiffres, il est remarquable qu'ils ne soient pas influencés par la construction linguistique des mots-nombres irréguliers, comme par exemple dans l'item «  $89 = 4 + 20 + 9$  ». Un élément vient cependant nuancer cette interprétation : le fait que la présentation de l'item soit uniquement en code indo-arabe. Il est donc possible que l'enfant manipule ce code sans avoir à le transcoder en code oral.

L'utilisation de la VP est observable par les épreuves de lecture et de dictée. Très peu d'écriture littérale terme à terme (codé en type « syntaxique » dans la présente étude) ont été relevées, en comparaison des travaux de Byrge et al. (2014) (discuté en partie Limites). La majorité des erreurs de type lexical peut suggérer une nouvelle fois que les enfants sont aussi sensibles à la grandeur globale du nombre : les dizaines sont composées uniquement de deux chiffres, les centaines de trois chiffres, etc. Les erreurs syntaxiques sont surtout retrouvées chez les CE1 pour les items des nombres supérieurs à 99, confirmant ainsi le modèle d'Herzog et al. (2019), proposant que la relation de regroupement unité-dizaine ne se généralise pas aux grandeurs supérieures sans repasser par un apprentissage explicite pour vérifier le regroupement à 10 : dix dizaines est égal à une centaine.

En conclusion, les compétences observées sont cohérentes avec les éléments exposés dans la recherche actuelle : le système linguistique influence peu les enfants au développement typique, les compétences procédurales de la VP semblent acquises avant les compétences conceptuelles, et la connaissance des valeurs de positions semble acquise tôt dans le processus de développement de la VP. Les résultats de la présente étude suggèrent que cette connaissance de valeur de position se développerait avant l'apport du vocabulaire mathématique. Contrairement à la littérature, cette présente étude observe peu d'erreurs terme à terme et suggère que les enfants pourraient être sensibles à la grandeur globale des nombres, c'est-à-dire au nombre de chiffres au sein du nombre. Cette sensibilité pourrait être une étape du développement de la compréhension de la VP.

Cependant, toutes ces interprétations issues des productions de la population étudiée présentement doivent être nuancées à l'égard de certaines limites inhérentes aux conditions de l'étude.

## **2 Limites de l'étude et perspectives de recherche**

Bien que les résultats soient encourageants, il est nécessaire de porter un regard critique sur certains points de l'étude. Tout d'abord, en raison de la faisabilité de l'étude, l'échantillon de participants est statistiquement mince. Ensuite, la population étudiée, scolarisée en milieu privé, est dans l'immense majorité issue d'un niveau socio-économique élevé, peu représentatif de l'entièreté des enfants français. De plus, le niveau socio-économique peut avoir une influence sur les apprentissages (OCDE, 2011). Cela pourrait expliquer pourquoi très peu d'erreurs littérales terme à terme ont été retrouvées dans les productions de la population étudiée, alors que Byrge et al. (2014) ont décrit cette erreur comme très fréquente chez les enfants en GSM. Aussi, la proportion d'enfants de GSM ayant réussi l'épreuve UDC doit être vérifiée sur un échantillon plus représentatif. Enfin, toujours en raison de la faisabilité de l'étude, la récolte des données s'est déroulée en décembre de l'année scolaire en cours, soit trois mois après le début de l'année scolaire. La passation des tests semble être arrivée trop tôt dans l'année pour être tout à fait représentative du niveau de la classe étudiée. En raison de ces trois limites, il serait pertinent de reconduire l'étude sur un échantillon plus important et composé d'enfants issus de niveaux socioéconomiques plus variés, incluant les milieux défavorisés. Enfin, une passation des tests entre avril et juin de l'année scolaire en cours serait plus en accord avec l'étalonnage de la batterie Examath 5-8. Dans ces conditions, il ne serait pas surprenant d'obtenir des résultats un peu différents. Cette prochaine étude permettra de former une norme fiable. L'inclusion d'enfants porteurs de Trouble des Apprentissages Mathématiques (DSM-5 ; 2015) sera nécessaire pour déterminer la sensibilité du module et ainsi garantir le pouvoir diagnostique de la batterie.

D'autres propriétés psychométriques restent à investiguer pour certifier la validité du module. Concernant la validité de surface, c'est-à-dire la recevabilité des utilisateurs, les épreuves ont été très appréciées des participants par son aspect ludique et numérique. Certains ont même souhaité refaire certaines tâches pour le plaisir, mais ces impressions doivent être objectivées statistiquement par un questionnaire de satisfaction par exemple. La validité de contenu, c'est-à-dire le lien

avec les modèles théoriques, semble bonne en regard de la revue de littérature présente. Les données expérimentales récoltées confortent les modèles théoriques actuels. L'avis d'experts pourrait également permettre d'estimer la validité de contenu. L'étude de la validité de critère devra être complétée par l'étude de la validité prédictive par une corrélation des scores aux tâches de la batterie avec les performances scolaires par exemple. Enfin, la validité de construit devra être complétée avec une étude de la validité factorielle.

Aussi, maintenant que la cohérence interne du module a été démontrée par le calcul de corrélations entre les épreuves, il serait intéressant d'affiner l'analyse en testant la cohérence inter-items. Cela permettrait d'appuyer les hypothèses dégagées par l'analyse des erreurs notamment pour l'épreuve de Décomposition d'addition pour laquelle les résultats par items ont été confrontés pour inférer quelques déductions.

Enfin, l'analyse des erreurs de transcodage aurait pu être plus approfondie avec un codage plus fin des productions que celui utilisé, classant les erreurs lexicales, syntaxiques ou mixtes. Questionner les enfants directement après leurs productions par une question large et ouverte pourrait aussi nous éclairer davantage sur la compréhension qu'ils ont du nombre et de sa construction.

### **3 Implications cliniques**

Ce présent travail, en testant la validité du module de numération d'Examath 5-8, s'inscrit dans le développement d'une nouvelle batterie d'évaluation. Les compétences en numération en début d'apprentissage et plus particulièrement la compréhension approfondie de la VP n'étant pas évaluables avec les outils existants, il s'agit d'une avancée considérable pour la profession orthophonique. Ainsi, les nouvelles tâches développées, inspirées de la recherche : décomposition d'addition, jugement d'association d'un nombre et d'une quantité analogique et enfin représentation de quantités avec des objets virtuels à partir de nombres indo-arabes, seront un apport conséquent à l'analyse orthophonique des compétences en numération d'un enfant en bilan. Aussi, la création de la norme critériée, même si elle ne peut être représentative de l'entièreté des enfants français de cette tranche d'âge, a offert de définir les compétences attendues aux classes de GSM, CP et CE1 en numération. Enfin, l'analyse des compétences numériques des enfants à chaque niveau scolaire a permis d'observer les compétences d'enfants au développement typique et de les confronter aux modèles théoriques actuels.

## V. Conclusion

Le développement de la compréhension de la VP est long et jalonné d'étapes, mais primordial à toute activité mathématique. C'est un défi pour tous les enfants en apprentissage. Trop peu d'outils dans la clinique orthophonique permettent d'évaluer finement la compréhension de la VP chez les enfants en apprentissage.

Les objectifs de ce mémoire ont été atteints : le module de numération de la nouvelle batterie d'évaluation Examath 5-8 possède une bonne cohérence interne, une bonne validité concomitante et une bonne validité de construit en lien avec les caractéristiques – c'est-à-dire le niveau scolaire – des individus. Les données obtenues ont permis l'élaboration d'une norme critériée, révélant la pertinence, aux niveaux scolaires étudiés, des nouvelles tâches de compréhension de VP proposées au sein de la batterie Examath 5-8. Ces nouvelles tâches, inspirées de la recherche, permettent d'analyser plus finement le stade de compréhension de l'enfant. L'analyse des productions des enfants a permis une confrontation des données expérimentales aux données développementales issues de la recherche. Certaines compétences observées confirment les modèles théoriques, d'autres observations permettent d'apporter quelques propositions. Ces éléments suggèrent une bonne validité de contenu du module de numération.

D'autres mesures psychométriques doivent être étudiées pour valider le module, comme la validité de surface, la validité prédictive et la sensibilité par exemple. En somme, la batterie Examath 5-8 améliorera considérablement l'évaluation des compétences en numération des enfants en apprentissage dans le cadre des bilans réalisés par les orthophonistes.

## VI. Bibliographie

- American Psychiatric Association. (2015). *DSM-5—Manuel diagnostique et statistique des troubles mentaux*. Elsevier Health Sciences.
- Aunio, P., & Räsänen, P. (2016). Core Numerical Skills for Learning Mathematics in Children Aged Five to Eight Years—A Working Model for Educators. *European Early Childhood Education Research Journal*, 24(5), 684-704.  
<https://doi.org/10.1080/1350293X.2014.996424>
- Bouchard, M.-E. G., Fitzpatrick, E. M., & Olds, J. (2009). Analyse psychométrique d'outils d'évaluation utilisés auprès des enfants francophones. *Revue canadienne d'orthophonie et d'audiologie*, 33(3), 129-139.
- Byrge, L., Smith, L. B., & Mix, K. S. (2014). Beginnings of place value : How preschoolers write three-digit numbers. *Child Development*, 85(2), 437-443.  
<https://doi.org/10.1111/cdev.12162>
- Cawley, J. F., Parmar, R. S., Lucas-Fusco, L. M., Kilian, J. D., & Foley, T. E. (2007). Place Value and Mathematics for Students with Mild Disabilities : Data and Suggested Practices. *Learning Disabilities: A Contemporary Journal*, 5(1), 21-39.
- Chan, B. M., & Ho, C. S. (2010). The cognitive profile of Chinese children with mathematics difficulties. *Journal of Experimental Child Psychology*, 107(3), 260-279. <https://doi.org/10.1016/j.jecp.2010.04.016>
- Chan, W. W. L. (2014). Understanding and processing numbers among Chinese children. *Psychology & Neuroscience*, 7(4), 583-591.  
<https://doi.org/10.3922/j.psns.2014.4.18>
- Chan, W. W. L., Au, T. K., Lau, N. T. T., & Tang, J. (2017). Counting errors as a window onto children's place-value concept. *Contemporary Educational*



- Psychology*, 51, 123-130. <https://doi.org/10.1016/j.cedpsych.2017.07.001>
- Chan, W. W. L., Au, T. K., & Tang, J. (2014). Strategic counting : A novel assessment of place-value understanding. *Learning and Instruction*, 29, 78-94.  
<https://doi.org/10.1016/j.learninstruc.2013.09.001>
- Chandler, C., C., & Kamii, C. (2009). Giving change when payment is made with a dime : The difficulty of tens and ones. *Journal for Research in Mathematics Education*, 40, 97-118.
- Cobb, P., & Wheatley, G. (1988). Children's Initial Understandings of Ten. *Focus on Learning Problems in Mathematics*, 10(3), 1-28.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). L. Erlbaum Associates.
- Dehaene, S. (2010). *La bosse des maths : Quinze ans après* (Nouv. édition revue et augmentée). O. Jacob.
- Desoete, A. (2015). Cognitive predictors of mathematical abilities and disabilities. In *The Oxford handbook of numerical cognition* (p. 915-932). Oxford University Press.
- Dietrich, J. F., Huber, S., Dackermann, T., Moeller, K., & Fischer, U. (2016). Place-value understanding in number line estimation predicts future arithmetic performance. *British Journal of Developmental Psychology*, 34(4), 502-517.  
<https://doi.org/10.1111/bjdp.12146>
- Dowker, A., & Roberts, M. (2015). Does the transparency of the counting system affect children's numerical abilities? *Frontiers in Psychology*, 6.  
<https://doi.org/10.3389/fpsyg.2015.00945>
- Fosnot, C., Towmey, & Dolk, M. (2001). *Young Mathematicians at Work : Constructing number sense, addition, and subtraction*. Heinemann.

- Fritz, A., & Ricken, G. (2008). *Rechenschwäche*. Reinhardt.
- Fuson, K. C., Smith, S. T., & Cicero, A. M. L. (1997). Supporting Latino First Graders' Ten-Structured Thinking in Urban Classrooms. *Journal for Research in Mathematics Education*, 28(6), 738. <https://doi.org/10.2307/749640>
- Fuson, K. C., Wearne, D., Hiebert, J. C., Murray, H. G., Human, P. G., Olivier, A. I., Carpenter, T. P., & Fennema, E. (1997). Children's conceptual structures for multidigit numbers and methods of multidigit addition and subtraction. *Journal for Research in Mathematics Education*, 28(2), 130-162. <https://doi.org/10.2307/749759>
- Gelman, R., & Meck, E. (1983). Preschoolers' counting : Principles before skill. *Cognition*, 13(3), 343-359. [https://doi.org/10.1016/0010-0277\(83\)90014-8](https://doi.org/10.1016/0010-0277(83)90014-8)
- Gervasoni, A. (2007). Assessing and teaching children who have difficulty learning arithmetic. *Educational & Child Psychology*, 24, 40-53.
- Glaser, R. (1963). Instructional technology and the measurement of learning outcomes : Some questions. *American Psychologist*, 18(8), 519-521. <https://doi.org/10.1037/h0049294>
- Hanich, L. B., Jordan, N. C., Kaplan, D., & Dick, J. (2001). Performance across different areas of mathematical cognition in children with learning difficulties. *Journal of Educational Psychology*, 93(3), 615-626. <https://doi.org/10.1037/0022-0663.93.3.615>
- Helloin, M.-C., & Lafay, A. (en préparation). *Examath 5-8, batterie informatisée d'examen des habiletés mathématiques pour les enfants de 5 à 8 ans*. HappyNeuron.
- Herzog, M., Ehlert, A., & Fritz, A. (2019). Development of a Sustainable Place Value Understanding. In Fritz A., Haase V., Räsänen P. (eds) *International*

*Handbook of Mathematical Learning Difficulties* (p. 561-579). Springer, Cham

[https://doi.org/10.1007/978-3-319-97148-3\\_33](https://doi.org/10.1007/978-3-319-97148-3_33)

Ivanova, M. V., & Hallowell, B. (2013). A tutorial on aphasia test development in any language : Key substantive and psychometric considerations. *Aphasiology*, 27(8), 891-920. <https://doi.org/10.1080/02687038.2013.805728>

Kamii, C. (1986). Place value : An explanation of its difficulty and educational implications for the primary grades. *Journal of Research in Childhood Education*, 1(2), 75-86. <https://doi.org/10.1080/02568548609594909>

Koppel, H. (1998). *Difficultés en mathématiques, évaluation et rééducation*. Papyrus.

Ladel, S., & Kortenkamp, U. (2016). Development of a Flexible Understanding of Place Value. In T. Meaney, O. Helenius, M. L. Johansson, T. Lange, & A. Wernberg (Éds.), *Mathematics Education in the Early Years : Results from the POEM2 Conference, 2014* (p. 289-307). Springer International Publishing. [https://doi.org/10.1007/978-3-319-23935-4\\_16](https://doi.org/10.1007/978-3-319-23935-4_16)

Lafay, A., Adrien, E. & Osana, H. P. (2020, Apr 17 - 21) *Second Graders' Domain-Specific Conceptual Difficulties in Base 10 Numeration* [Paper Session]. AERA Annual Meeting San Francisco, CA <http://tinyurl.com/u7n4h4y> (Conference Canceled - coronavirus)

Lafay, A., & Cattini, J. (2018). Analyse psychométrique des outils d'évaluation mathématique utilisés auprès des enfants francophones. *Revue canadienne d'orthophonie et d'audiologie*, 42, 127-144.

Laski, E. V., Schiffman, J., Shen, C., & Vasilyeva, M. (2016). Kindergartners' base-10 knowledge predicts arithmetic accuracy concurrently and longitudinally. *Learning and Individual Differences*, 50, 234-239.

<https://doi.org/10.1016/j.lindif.2016.08.004>

- Leclercq, A.-L., & Veys, E. (2014). Réflexions sur le choix de tests standardisés lors du diagnostic de dysphasie. *Approche Neuropsychologique des Apprentissages chez l'Enfant*, 26, 374-382.
- MacDonald, A. (2008). « But what about the oneths? » A Year 7 student's misconception about decimal place value. *Australian Mathematics Teacher*, 64(4), 12-16.
- Mann, A., Moeller, K., Pixner, S., Kaufmann, L., & Nuerk, H.-C. (2012). On the development of Arabic three-digit number processing in primary school children. *Journal of Experimental Child Psychology*, 113(4), 594-601.  
<https://doi.org/10.1016/j.jecp.2012.08.002>
- Mark, W., & Dowker, A. (2015). Linguistic influence on mathematical development is specific rather than pervasive : Revisiting the Chinese Number Advantage in Chinese and English children. *Frontiers in Psychology*, 6.  
<https://doi.org/10.3389/fpsyg.2015.00203>
- McGuire, P., & Kinzie, M. B. (2013). Analysis of place value instruction and development in pre-kindergarten mathematics. *Early Childhood Education Journal*, 41(5), 355-364. <https://doi.org/10.1007/s10643-013-0580-y>
- Métral, E. (2008). *Malette B-LM Cycle 2*. Orthopratic.
- Mix, K. S., Smith, L. B., Stockton, J. D., Cheng, Y.-L., & Barterian, J. A. (2017). Grounding the symbols for place value : Evidence from training and long-term exposure to base-10 models. *Journal of Cognition and Development*, 18(1), 129-151. <https://doi.org/10.1080/15248372.2016.1180296>
- Moeller, K., Pixner, S., Zuber, J., Kaufmann, L., & Nuerk, H. C. (2011). Early Place-Value Understanding as a Precursor for Later Arithmetic Performance—A Longitudinal Study on Numerical Development. *Research in Developmental*

*Disabilities: A Multidisciplinary Journal*, 32(5), 1837-1851.

<https://doi.org/10.1016/j.ridd.2011.03.012>

Nuerk, H.-C., Moeller, K., & Willmes, K. (2015). Multi-digit number processing : Overview, conceptual clarifications, and language influences. In R. C. Kadosh & A. Dowker (Éds.), *The Oxford handbook of numerical cognition*. (2015-44582-007; p. 106-139). Oxford University Press.

Nunnally, J. C. (1978). *Psychometric Theory*. McGraw-Hill.

OCDE. (2011). « Le milieu des élèves affecte-t-il leur performance ? ». In *Regards sur l'éducation 2011 : Les indicateurs de l'OCDE* (Editions OCDE). Paris.

<https://doi.org/10.1787/eag-2011-9-fr>.

Osana, H. P., Lafay, A., & Blondin, A. (2018). *PicPVT, une tâche d'évaluation de la compréhension de la double représentation des symboles dans les nombres : Validation chez les enfants de 1e année de primaire*. 2. Acte de congrès, rencontres d'Espace Mathématique Francophone, Paris.

Pixner, S., Moeller, K., Hermanova, V., Nuerk, H.-C., & Kaufmann, L. (2011). Whorf reloaded : Language effects on nonverbal number processing in first grade—A trilingual study. *Journal of Experimental Child Psychology*, 108(2), 371-382.

<https://doi.org/10.1016/j.jecp.2010.09.002>

Plante, E., & Vance, R. (1994). Selection of preschool language tests : A data-based approach. *Language, Speech, and Hearing Services in Schools*, 25(1), 15-24.

<https://doi.org/10.1044/0161-1461.2501.15>

Ross, S. H. (1989). Parts, Wholes, and Place Value : A Developmental View. *The Arithmetic Teacher*, 36(6), 47-51.

Van de Walle, J. A., Karp, K. S., & Bay-Williams, J. M. (2004). *Elementary and middle school mathematics : Teaching developmentally*. Pearson.

Van Nieuwenhoven, C., Grégoire, J., & Noël, M.-P. (2001). *Tedi-math : Test diagnostique des compétences de base en mathématiques*. Editions du Centre de Psychologie appliquée.

Wainer, H., & Braun, H. I. (1988). *Test Validity*. Lawrence Erlbaum Associates, Inc.

## VII. Annexes

### Annexe A : Illustration des résultats de l'hypothèse 3 – Validité de construit en lien avec le niveau scolaire

