ÉVALUER LA RÉSOLUTION DE PROBLÈMES ADDITIFS AU CP : UNE ÉTUDE EXPLORATOIRE AUTOUR DE LA TAILLE DES NOMBRES

Nadine GRAPIN[1]
Univ Paris Est Créteil, Université ParisCité, CY Cergy Paris Université, Univ.Lille, Univ Rouen, LDAR, F-94010 Créteil, France

Éric MOUNIER[2]
Univ Paris Est Créteil, Université ParisCité, CY Cergy Paris Université, Univ.Lille, Univ Rouen, LDAR, F-94010 Créteil, France

Résumé. Nous proposons dans cet article une étude exploratoire autour de l'évaluation des connaissances des élèves en résolution de problèmes arithmétiques additifs verbaux basiques. En jouant à la fois sur différentes classes de problèmes et sur la taille des nombres, nous obtenons des résultats quantitatifs et qualitatifs sur les performances d'élèves de CP, basés sur une interprétation de leurs réponses au regard du schéma de Verschaffel et De Corte (2008) modélisant l'activité en résolution de problèmes. Nous proposons alors des pistes pour l'enseignant et pour le chercheur afin d'évaluer les connaissances des élèves lors de la résolution de problèmes.

Mots-clés. Résolution de problèmes arithmétiques, évaluation, taille des nombres, cours préparatoire (CP).

Introduction

La dernière enquête PRAESCO (DEPP, 2021) a montré que les pratiques des enseignants de CM2 (grade 5, élèves âgés de 10-11 ans) en résolution de problèmes étaient assez variées, puisque la moitié d'entre eux *« propose souvent des problèmes pour découvrir une notion (49 %), pour apprendre à chercher (50 %) ou pour se confronter à la complexité (problèmes à étapes sans question intermédiaire, 46 %) »*. Si on s'intéresse plus spécifiquement aux problèmes arithmétiques verbaux, c'est-à-dire aux problèmes qui *« racontent des histoires […] qui sont donnés avec des mots et font intervenir peu de symbolisme mathématique »* (Feyfant, 2015, p. 9), on observe que quel que soit le niveau scolaire considéré, les élèves français rencontrent des difficultés, notamment dans la résolution de problèmes arithmétiques basiques, *« à deux données […] où il s'agit de déterminer une troisième valeur […], à énoncé court, syntaxe simple, sans information superflue »* (Houdement, 2017) ; ce constat reste encore d'actualité, comme le confirment les résultats aux évaluations nationales au CP et au CE1 (DEPP, 2023). Même si on ne peut pas considérer ces résultats sans interroger les problèmes proposés, les conditions de passation, la détermination des seuils, etc., ce constat rejoint cependant celui des enseignants de CP et CE1 d'une école de Montreuil (dans le département de Seine-Saint-Denis en France) avec lesquels nous menons une recherche collaborative depuis 2019 (Blanchouin *et al.*, 2022a).

[1] nadine.grapin@u-pec.fr

[2] eric.mounier@u-pec.fr

Par ailleurs, depuis une vingtaine d'années en France, différentes directives institutionnelles incitent les professeurs des écoles à pratiquer une évaluation *« constructive, qui régule les enseignements et soutient les apprentissages* » (MENJS, 2016, p. 1), mais elle reste encore peu présente dans les pratiques effectives des enseignants (Sayac, 2019). Alors que ceux-ci expriment leur besoin de formation pour mieux évaluer les connaissances de leurs élèves, mais aussi pour améliorer les compétences de ces derniers en mathématiques (DEPP, 2021), nous observons que, hormis des évaluations nationales à visée diagnostique proposées à l'entrée au CP, au CE1 et au CM1, peu d'outils sont mis à leur disposition par le Ministère de l'Éducation Nationale pour les orienter dans leurs pratiques d'évaluation en classe.

En tant que chercheurs et formateurs d'enseignants, nous avons eu alors comme objectif de concevoir une ingénierie didactique évaluative (Vantourout & Maury, 2017) c'est à dire un dispositif qui permettrait à l'enseignant d'identifier les connaissances de ses élèves *« pour ensuite proposer une régulation didactique de leur enseignement, en fonction de l'acquisition diagnostiquée des compétences et des apprentissages visés »* (Vantourout, 2020, p. 118). Nous avons choisi de nous centrer sur une modalité d'évaluation spécifique, en passation collective, en papier crayon, sans interaction directe avec l'enseignant ; ce qui correspond aux modalités d'évaluation formelle les plus souvent pratiquées à l'école élémentaire. Pour ce faire, et avant de proposer à l'enseignant un tel dispositif, il est nécessaire de le « tester » auprès d'élèves, dans des conditions ordinaires de classe (*ibid.*, p. 108). C'est pourquoi, dans notre recherche, nous tirons des résultats sur les performances des élèves en résolution de certains problèmes arithmétiques verbaux (RPAV) à partir d'une série de problèmes additifs que nous leur avons soumis. Dans cet article, nous discutons en particulier de l'effet de la taille des nombres sur la réussite et sur les procédures déployées pour résoudre ces problèmes.

Après avoir présenté des éléments théoriques autour de l'évaluation et de la résolution de problèmes, nous précisons la problématique et notre méthodologie. Les résultats de notre recherche sur les connaissances des élèves sont ensuite présentés et accompagnés de propositions à destination de l'enseignant pour qu'il puisse diagnostiquer plus précisément les connaissances et compétences de ses élèves et proposer des régulations adaptées.

1. Évaluer les connaissances et les compétences en résolution de problèmes arithmétiques : éléments théoriques

En nous référant à la synthèse de Mottier-Lopez (2015, p. 43), et comme de nombreux autres chercheurs, nous considérons cinq composantes participant au processus évaluatif :

> • *l'objet à évaluer et ses dimensions à analyser ;*
>
> • *les attentes vis-à-vis de cet objet [...] ;*
>
> • *le recueil d'un ensemble d'informations en rapport avec l'objet au regard des possibilités du réel ;*
>
> • *l'interprétation des informations recueillies [...] ;*
>
> • *la formulation d'une appréciation qui doit pouvoir être communiquée et fonder des prises de décision.*

Afin de pouvoir identifier les différentes dimensions de l'objet à évaluer et nous intéresser ensuite à la validité des tâches évaluatives, nous devons au préalable définir les compétences et connaissances en jeu dans l'activité[3] de résolution de problèmes.

1.1. Modélisation de l'activité de l'élève en résolution de problèmes

Dans le but de modéliser l'activité de l'élève en résolution de problèmes, nous nous sommes appuyés sur le modèle de Verschaffel et de Corte (2008) que nous avons réinterprété (*cf.* figure 1) pour prendre en compte la spécificité des problèmes proposés et l'âge des élèves (Chenevotot *et al.*, 2022). En effet, nous nous limitons dans le cadre de notre recherche à des problèmes arithmétiques verbaux basiques, qui ont déjà fait l'objet d'un enseignement et pour lesquels le modèle mathématique est *a priori* connu des élèves : il peut s'agir d'un modèle additif ou soustractif pouvant conduire à l'écriture d'un calcul (addition, soustraction ou addition à trou) ou à la représentation d'une ou des deux collections (par un dessin plus ou moins figuratif ou à l'aide de matériel : jetons, cubes, etc.).

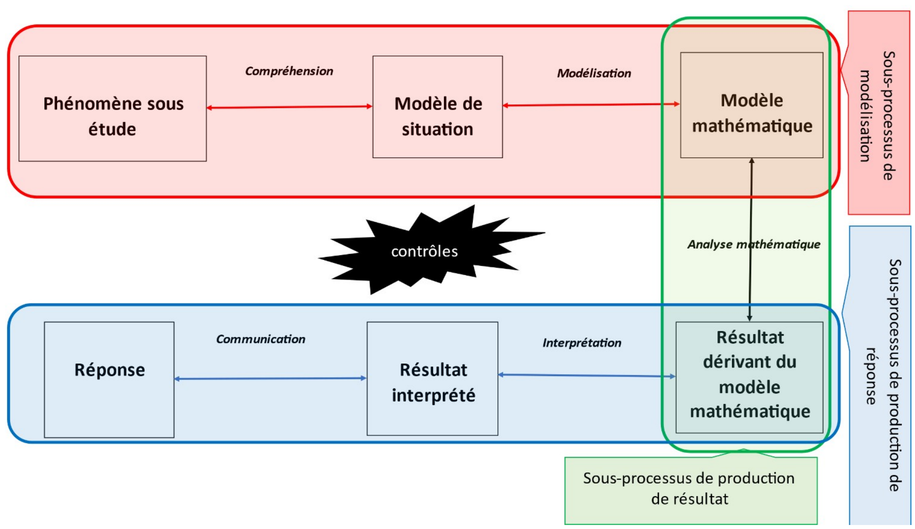


Figure 1 *: Trois sous-processus pour interpréter le schéma de Verschaffel et de Corte (2008) dans le cadre de la RPAV.*

Nous définissons ainsi trois sous-processus (un de modélisation, un de production de résultat et un de production de réponse) qui nous permettent de rendre compte d'un point de vue théorique de l'activité de résolution de problèmes arithmétiques avec, en parallèle, sur l'ensemble de ces sous-processus, une activité possible de contrôle. La résolution de problèmes ne se résume pas à un processus linéaire composé de ces trois sous-processus puisque des allers-retours sont possibles, en particulier si l'activité de contrôle conduite par l'élève l'amène à constater une erreur sur le produit d'un des sous processus ; la résolution de problèmes ne se limite pas non plus à écrire une opération, faire un calcul, l'effectuer et écrire une phrase réponse, le processus est bien plus complexe et par conséquent il est d'autant plus compliqué d'évaluer les connaissances des élèves impliquées dans la résolution.

En ce qui concerne la résolution de problèmes arithmétiques verbaux mettant en jeu des

[3] Nous entendons par activité, *« ce que développe le sujet lors de la réalisation de la tâche : non seulement ses actes extériorisés, mais aussi des inférences, les hypothèses qu'il fait, les décisions qu'il prend, la manière dont il gère son temps, mais aussi son état personnel »* (Rogalski, 2003, pp. 349-350).

quantités discrètes, et plus particulièrement pour certaines classes de problèmes relevant de la structure additive (Vergnaud, 1991 ; Riley *et al.*, 1984), un des enjeux du CP (MENJS, 2021) est d'amener les élèves à faire évoluer leurs stratégies initiales de résolution basées sur du dénombrement (comptage, sur-comptage, décomptage) à des stratégies s'appuyant sur du calcul : c'est-à-dire opérer sur les nombres et non plus sur des collections (Conne, 1989). Ceci nécessite une reconnaissance de l'opération en jeu et l'écriture d'un calcul, qui est à relier au sous-processus de modélisation, puis l'effectuation du calcul mentalement ou de façon posée. Or, pour des élèves de cet âge, ce qu'ils peuvent donner à voir du modèle mathématique peut être sous forme d'une collection (de doigts, d'objets dessinés plus ou moins figuratifs) qui leur sert également de support pour opérer directement sur les quantités (par comptage, sur-dé comptage) : les sous-processus de modélisation et de production de résultat sont ainsi parfois difficiles à dissocier et à distinguer. Par exemple, pour résoudre le problème « Alicia a 5 billes ; elle en gagne 4 à la récréation. Combien en a-t-elle à la fin de la récréation ? », l'élève peut lever cinq doigts puis quatre autres pour à la fois modéliser la situation et produire le résultat (9) (Grapin *et al.*, 2022).

S'il est évident qu'un manque de connaissances en calcul peut empêcher d'obtenir un résultat correct lors du sous-processus de production de résultat malgré un modèle mathématique pertinent, la taille des nombres peut également empêcher de se forger un modèle de situation adapté (que ce soit mentalement ou via des dessins). Par exemple, dans le problème suivant de réunion avec recherche du tout : « Jean a 45 billes vertes et rouges ; il a 18 billes vertes, les autres sont rouges. Combien a-t-il de billes rouges ? ». Les réponses 33, 63 et 62 (que pourraient produire des élèves) sont toutes trois erronées, mais la réponse 33 pourrait correspondre plutôt à une erreur de calcul de $45-18$ alors que 63 correspond à la somme de 45 et 18. 62 peut quant à lui correspondre à une erreur d'obtention de la somme de 45 et 18 (que ce soit par un calcul ou un dénombrement). Si on reprend le problème avec des petits nombres, par exemple 6 billes vertes et rouges dont 2 billes seulement sont vertes, la situation change encore. On peut s'attendre essentiellement à deux réponses : 8 et 4. Cela met alors en évidence qu'il existe des modèles de situation ayant souvent leur origine dans la manipulation de matériel, qui ne sont pas liés à des connaissances en calcul.

1.2. Validité d'un test

Classiquement, s'assurer qu'un test est valide, c'est s'assurer qu'il évalue bien ce qu'il doit évaluer (Vantourout, 2020, p. 112). Les travaux menés en didactique des mathématiques ces dix dernières années (Grapin & Grugeon-Allys, 2018 ; Grapin *et al.*, 2022) ont conduit à définir une approche didactique de la validité prenant en compte à la fois l'épistémologie des savoirs en jeu dans les tâches évaluatives et l'activité de réponse de l'élève. Pour ces auteurs, s'assurer de la validité didactique d'un test (une tâche ou un ensemble de tâches) passe par une analyse *a priori* des tâches du test au regard d'une référence épistémologique ; elle conduit à définir, pour chaque tâche, en fonction des valeurs des principales variables didactiques, les procédures de résolution possibles et la complexité de la tâche. Ce premier niveau d'analyse, localement, tâche par tâche, permet de s'assurer que les connaissances nécessaires à la résolution de la tâche sont celles qui sont évaluées, de tenir compte des spécificités des tâches et de ce qu'elles pourraient engendrer dans l'activité de l'élève. Un deuxième niveau d'analyse concerne l'ensemble des tâches du test. Il s'agit de s'assurer de la couverture du domaine évalué, sans manque ni redondance mais aussi d'anticiper certaines mises en relation et hiérarchies selon la complexité des tâches ; ce que nous illustrerons à partir de la série de problèmes que nous proposons.

Une telle approche, dans laquelle l'analyse *a priori* des tâches joue un rôle central, permet de concevoir le test en sélectionnant les types de tâches et en choisissant les valeurs de variables selon les connaissances que l'on souhaite évaluer. Elle permet également d'interpréter les réponses proposées en les reliant éventuellement à des procédures erronées qui ont été identifiées lors de l'analyse *a priori*. Enfin, si on se place à un niveau plus global, les réponses données par un même élève à des tâches différentes peuvent être mises en relation et révéler la disponibilité de ses connaissances ou la résistance de certaines erreurs. Par exemple, en résolution de problèmes, des démarches superficielles (Van Dooren *et al.*, 2015 ; Fagnant, 2018) s'appuyant uniquement sur les mots inducteurs (comme associer systématiquement le terme « ajouter » dans le texte du problème à une addition à faire), peuvent être repérées en mettant en relation les réponses d'un même élève à plusieurs problèmes. Nous allons utiliser le terme de « démarche superficielle » comme caractérisant une activité contribuant essentiellement au sous-processus de production de résultat, sans qu'il n'y ait eu de modèle de situation ni de résultat interprété. S'il y a eu des contrôles, ils n'ont concerné que les nombres (dénombrements, calculs). En particulier il n'y a pas eu de qualification au sens de Houdement (2017). Dans le cas où une activité participant au sous-processus de modélisation a été déployée nous ne parlons pas de démarche superficielle, même si cette activité n'a pas permis de mobiliser un modèle adapté (Mounier *et al.*, à paraître).

2. Problématique

Le modèle que nous avons choisi pour étudier l'activité de l'élève témoigne de la complexité de cette activité.

Même si ces processus sont dépendants les uns des autres, qu'il peut y avoir des allers-retours entre ceux-ci, évaluer les connaissances et compétences des élèves en résolution de problèmes demande de pouvoir recueillir des informations sur chacun des sous-processus et sur les contrôles effectués par l'élève, d'interpréter ces informations puis de prendre une décision (qui dépend ici des fonctions et des objectifs assignés à l'évaluation).

Accompagner l'enseignant dans l'évaluation des connaissances de ses élèves implique donc de lui proposer des tâches demandant, pour être résolues, de mobiliser les connaissances qu'il souhaite évaluer, mais aussi de penser l'usage qui peut en être fait en classe ; pour reprendre Vantourout (2020, p. 118), *« une épreuve, un test, n'est pas valide en soi, mais dépend de l'usage que l'on en fait »*.

Il s'agit donc de sensibiliser l'enseignant au choix des problèmes arithmétiques qu'il propose mais aussi à l'interprétation qu'il fait de la réponse de l'élève. Quels problèmes proposer aux élèves pour qu'ils constituent un ensemble didactiquement valide (Grapin & Grugeon-Allys, 2018 ; Grapin & Mounier, 2019) ?

Pour poursuivre avec Vantourout (2020, p. 118) :

> *Avec les ingénieries évaluatives, les enseignants disposeraient de résultats quantitatifs (hiérarchies de réussite des problèmes par les élèves) et qualitatifs (procédures de résolution). Ils peuvent ainsi situer leurs élèves, sachant que les populations testées par les chercheurs ne constituent pas des échantillons représentatifs. On retrouve ainsi une « combinaison » qui repose à la fois sur le report à une norme et le rapport au développement formatif.*

Le développement d'une ingénierie évaluative nous conduit alors à produire des résultats quantitatifs et qualitatifs à destination de l'enseignant. Au-delà de participer au développement de l'ingénierie, ces résultats permettent également d'enrichir ceux scientifiques produits autour de la résolution de problèmes.

Beaucoup de paramètres, facteurs, éléments, variables entrent en jeu dans le choix des problèmes à résoudre et des modalités de passation ; et les élèves empruntent des cheminements cognitifs divers d'autant qu'ils sont jeunes (Mounier, 2010). Dans cette étude exploratoire, nous avons choisi de proposer des problèmes issus de classes de problèmes différentes (Vergnaud, 1991) en jouant sur la variable « taille des nombres ». En effet, si Butlen (2007) et Fayol (2008) ont souligné le lien qui peut exister entre les connaissances des élèves sur les nombres et les opérations et leurs performances dans la résolution de problèmes, il nous semble intéressant de mener une étude complémentaire ciblée sur la taille des nombres et dans une moindre mesure sur les propriétés arithmétiques des nombres et des opérations.

Nous souhaitons ainsi mettre l'accent sur le rôle que peuvent jouer les nombres (qui indiquent ici toujours des quantités discrètes) dans l'activité en RPAV. Plus précisément, nous voulons repérer les élèves pour lesquels la variable « taille des nombres » sera pertinente, c'est-à-dire aura un effet sur la réussite d'un « même » problème (même classe de problèmes, même contexte, même structure sémantique).

Des activités de contrôle peuvent intervenir dans une résolution de problèmes, y compris dans le sous-processus de production de réponse à partir d'un résultat numérique obtenu. Cependant, il faudrait observer finement l'activité de l'élève durant sa recherche ou mener des entretiens avec lui pour avoir des informations sur ces contrôles. Or nous voulons proposer un test qui permette une évaluation des connaissances des élèves à partir des réponses données. Notre test a donc pour objectif de présenter à l'enseignant, en fonction de la variable « taille des nombres » et de la classe de problèmes, des hiérarchies de réussite pour la résolution de certains problèmes arithmétiques (en termes de réussite/échec) et des démarches de résolution prenant en compte les sous-processus de modélisation et de production de résultats.

Nous explicitons dans la partie expérimentation nos choix pour concevoir le test selon des critères de validité didactique puis nous montrons en quoi les données produites permettent de répondre à nos objectifs.

3. Expérimentation

L'expérimentation a été menée en France en juin 2017 auprès de 83 élèves de CP issus de 6 classes différentes, toutes situées en REP[4] et dans lesquelles les enseignants utilisaient la même ressource : *Mon Année de Maths CP* (Mazollier *et al.*, 2016). Elle est composée d'un fichier pour l'élève, comportant le savoir à retenir et des exercices, ainsi que d'un guide pour l'enseignant pour mener en particulier les séances de découverte.

3.1. Problèmes retenus

Afin de concevoir un test valide au regard des objectifs décrits ci-avant, nous avons choisi des classes de problèmes (Vergnaud, 1991) conformes au programme scolaire (MENJS, 2018) et à celles présentes dans le manuel, à savoir : composition de mesures avec recherche du tout ou d'une des parties, transformation d'état (transformation positive ou négative) avec recherche de l'état final ou de la transformation (*cf.* annexe 1).

[4] Réseau d'Éducation Prioritaire. Les écoles de ces réseaux bénéficient de moyens supplémentaires, comme une formation plus importante des enseignants.

Pour chacune d'elles, nous avons proposé deux problèmes : l'un mettant en jeu des valeurs numériques inférieures à 10 et dont la somme est inférieure à 10 (« petits nombres » - PN) et l'autre avec au moins une des valeurs supérieures à 20 (« grands nombres » - GN). Par exemple, pour la classe de problèmes *Recherche de la transformation (transformation positive)*, le problème avec PN est celui proposé ci-dessous (*cf.* figure 2). Celui avec GN est identique, seules les données numériques changent : il y a 23 jetons dans la boîte au début et 53 jetons à la fin.

On cherche le nombre de jetons que l'on verse dans la boîte.

Au début, il y a 7 jetons dans la boîte.
Je verse tous les jetons du sachet dans la boîte.
A la fin, dans la boîte, il y a 10 jetons.

Quel est le nombre de jetons que j'ai versé dans la boîte ?
Écris ta réponse sur les pointillés.

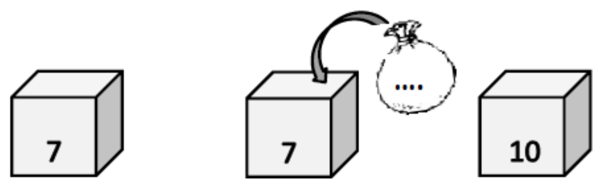


Figure 2 *: Enoncé du problème avec petits nombres pour la classe de problèmes* Recherche de la transformation (transformation positive).

Les douze problèmes proposés à chacun des 83 élèves sont récapitulés dans le tableau 1 ci-après et sont donnés en intégralité en annexe 1.

Classe de problèmes		Nombres en jeu	
		Petits Nombres (PN)	Grands Nombres (GN)
Réunion de deux parties (P_1 et P_2) avec recherche du tout.		$P_1=7$; $P_2=8$	$P_1=35$; $P_2=10$
Réunion avec recherche d'une des parties (P_2) connaissant une partie (P_1) et le tout.		$P_1=2$; $Tout=10$	$P_1=30$; $Tout=52$
Transformation (t) avec recherche de l'état final (E_f) connaissant l'état initial (E_i).	Ajout	$E_i=5$; $t=+3$	$E_i=13$; $t=+16$
	Retrait	$E_i=6$; $t=-2$	$E_i=37$; $t=-10$
Transformation avec recherche de la transformation (t) connaissant l'état initial (E_i) et l'état final (E_i).	Ajout	$E_i=7$; $E_f=10$	$E_i=23$; $E_f=53$
	Retrait	$E_i=9$; $E_f=5$	$E_i=46$; $E_f=26$

Tableau 1 *: Présentation synthétique des problèmes (classes de problèmes et nombres en jeu).*

Nous nous sommes assurés que la résolution demandait la mobilisation des connaissances que nous souhaitions évaluer et que, *a priori*, les informations recueillies permettraient d'inférer ces connaissances. Pour ce faire, nous avons d'abord choisi des problèmes qui sont tous issus d'une même situation de référence, celle dite de « la boîte » : dans une boîte opaque, on réunit deux collections de jetons ou on ajoute ou on retire un certain nombre de jetons. En proposant un contexte proche de celui du quotidien scolaire des élèves, nous avons souhaité éviter des effets de contexte sur la compréhension de la situation. En conservant le même pour l'ensemble du test, il était alors possible d'interpréter la différence de réussite à une même classe de problèmes, selon la seule variable qui avait une valeur différente, à savoir la taille des nombres. Les « grands nombres » ont été ensuite choisis afin que le calcul en lui-même (addition ou soustraction), s'il était envisagé par l'élève, ne présente pas de difficulté *a priori* : l'élève peut effectuer les calculs mentalement et s'il envisage de les poser, il n'y a pas de retenue.

La formulation de l'énoncé sous forme de bande dessinée en trois étapes et l'utilisation de flèches dans les illustrations s'appuient sur un symbolisme connu des élèves parce que similaire à celui proposé dans le fichier de leur manuel. La première phrase de l'énoncé indique à l'élève ce qu'il doit chercher. Toutes les phrases sont formulées avec des verbes conjugués au présent ; les indicateurs temporels dans les problèmes de transformation d'état sont indiqués par « au début » et « à la fin ». Chaque énoncé est accompagné d'une partie en quadrillage *seyes* et d'une phrase réponse à compléter.

3.2. Déroulement de la passation

Afin de concevoir une évaluation pouvant être proposée collectivement en classe, nous avons cherché, dès cette étude exploratoire, à nous mettre dans des conditions proches de celles qui existent habituellement en classe de CP lors de la passation d'évaluations formelles (passation collective en classe entière, énoncé écrit, réponse individuelle des élèves sur une feuille). Trois passations ont eu lieu trois jours de suite dans chacune des 6 classes de CP. Chaque élève disposait d'un livret composé de quatre problèmes et afin d'éviter une certaine pression évaluative, ces exercices ont été présentés comme des défis et non comme une suite d'évaluations.

Chaque problème a été lu une seule fois par le chercheur ; en parallèle de la lecture, les actions sur le matériel (jetons/boîte) ont été réalisées devant la classe et il a été indiqué à partir d'un agrandissement au tableau de l'énoncé, pour chaque action, ce qui correspondait sur le schéma. Un temps de réponse approximatif de trois minutes a été prévu mais il a été adapté selon l'avancée des élèves ; aucune aide ne leur a été apportée.

Chaque livret a été ensuite relevé et les réponses des élèves ont été reportées par les chercheurs dans une feuille de calcul (tableur) pour analyse.

4. Résultats

Après avoir présenté des résultats quantitatifs en termes de réussite-erreur sur l'ensemble des problèmes, nous montrons deux types d'analyses qualitatives complémentaires basées uniquement sur les réponses numériques des 83 élèves, l'une pour une classe de problèmes et l'autre en croisant les réponses aux problèmes avec PN et GN.

4.1. Résultats quantitatifs : analyse réussite / échec par problème

Un premier niveau d'analyse est relatif à l'exactitude de la réponse numérique donnée et conduit à un pourcentage de réussite (*cf.* tableau 2).

Classe de problèmes		Petits Nombres (PN)		Grands Nombres (GN)	
Réunion de deux parties (P_1 et P_2) avec recherche du tout.		$P_1=7$ $P_2=8$	$R:75\%$	$P_1=35$ $P_2=10$	$R:70\%$
Réunion avec recherche d'une des parties (P_2) connaissant une partie (P_1) et le tout.		$P_1=2$ $Tout=10$	$R:65\%$	$P_1=30$; $Tout=52$	$R:40\%$
Transformation (t) avec recherche de l'état final (E_f) connaissant l'état initial (E_i).	Ajout	$E_i=5$ $t=+3$	$R:66\%$	$E_i=13$ $t=+16$	$R:50\%$
	Retrait	$E_i=6$ $t=-2$	$R:60\%$	$E_i=37$ $t=-10$	$R:41\%$
Transformation avec recherche de la transformation (t) connaissant l'état initial (E_i) et l'état final (E_i).	Ajout	$E_i=7$ $E_f=10$	$R:65\%$	$E_i=23$ $E_f=53$	$R:29\%$
	Retrait	$E_i=9$ $E_f=5$	$R:52\%$	$E_i=46$ $E_f=26$	$R:35\%$

Tableau 2 *: Pourcentage de réussite (R) selon les problèmes.*

Nous retrouvons, comme dans beaucoup d'autres recherches antérieures (Fagnant, 2005, 2013 ; synthèse de Verschaffel & De Corte, 1997), une hiérarchie en termes de réussite selon les classes de problèmes, les problèmes de réunion avec recherche du tout et de recherche de l'état final avec transformation positive étant les mieux réussis ; la recherche de la transformation étant plus échouée, en particulier avec des grands nombres. L'effet « taille des nombres » est aussi visible, les problèmes avec petits nombres étant systématiquement mieux réussis que ceux avec grands nombres.

S'il est intéressant de rappeler ou de montrer à l'enseignant cette hiérarchie de réussite selon les classes de problèmes et la taille des nombres, se soucier uniquement de l'exactitude de la réponse ne permet pas de déterminer les démarches mobilisées par les élèves ni d'évaluer les connaissances dont ils font preuve ou non. Comment exploiter de telles données pout informer l'enseignant sur les démarches utilisées par chacun de ces élèves et les sous-processus en jeu dans la résolution ?

4.2. Exploitation de l'analyse *a priori* pour interpréter les réponses : illustration à l'aide d'un exemple de classe de problèmes

La passation ayant été réalisée en classe collectivement, les seuls éléments sur lesquels nous pouvons nous appuyer pour interpréter la réponse de l'élève, autre que la réponse elle-même, sont les traces écrites laissées lors de la résolution[5]. Or 88 % des productions en moyenne (95 % pour les problèmes avec des nombres inférieurs à 10, 81 % pour les autres) n'ont pas de traces apparentes autres que l'écriture chiffrée de la réponse. Savoir si l'élève a produit un modèle de situation pertinent, si le modèle choisi est adapté ou non et/ou s'il a commis une erreur de calcul reposerait donc dans la plupart des cas uniquement sur l'interprétation de ce nombre donné en réponse.

[5] Une expérimentation complémentaire a été menée par la suite en complétant l'analyse des productions écrites par des entretiens entre chercheurs et élèves (*cf.* Grapin *et al.*, 2022 ; Chenevotot *et al.*, 2022 ; Mounier *et al.*, à paraître).

Éléments d'analyse **a priori**

Étudions, à partir des deux problèmes de transformation négative avec recherche de l'état final, (figure 3) la façon dont l'analyse *a priori* permet de fournir une telle interprétation.

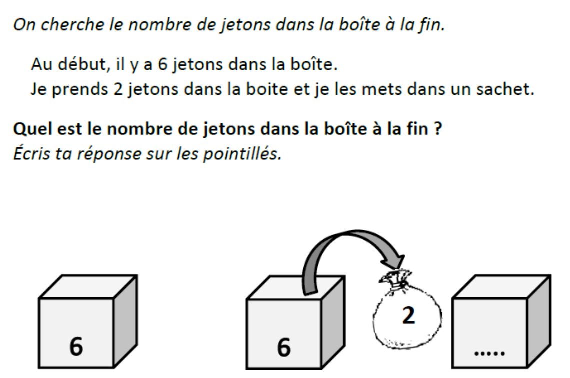
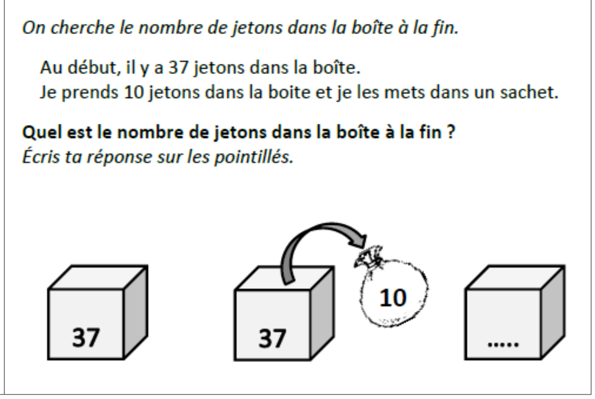
<table>
<tr><td>Transformation négative avec recherche de l'état final - « petits nombres ».</td><td>Transformation négative avec recherche de l'état final - « grands nombres ».</td></tr>
<tr><td>
</td><td>
</td></tr>
</table>

Figure 3 *: Enoncés des deux problèmes de transformation négative avec recherche de l'état final.*

Dans ces deux problèmes, sauf en cas de démarche superficielle, c'est à dire sans activité dans le sous-processus de modélisation ni de production de réponse, les réponses exactes (4 et 27) pourraient être interprétées comme une modélisation correcte du problème alliée à des connaissances avérées en dénombrement ou en calcul. Les réponses incorrectes à 1 près, mais aussi plus largement, celles inférieures à la différence des deux nombres en jeu dans le problème (réponses inférieures à 4 et inférieures à 27) pourraient être interprétées comme provenant essentiellement d'erreurs de calcul ou de dénombrement. S'il semble raisonnable de retenir cette hypothèse pour les réponses à 1 près, il y aurait nécessité de faire d'autres investigations pour les autres (en demandant par exemple des explications à l'élève).

D'autres types de réponses pourraient être interprétés comme une modélisation incorrecte du problème ou une absence d'activité dans le sous-processus de modélisation, qui pourrait alors révéler dans ce dernier cas uniquement une démarche superficielle. Par exemple 6+2 est convoqué par l'image de 6 et 2 présents dans la bande dessinée, sans tenir compte de l'énoncé du problème ni des actions matérielles effectuées par le chercheur devant la classe. Ainsi les réponses 8 et 47 (correspondant à la somme des deux nombres en jeu dans le problème) ou encore 14 et 84 (correspondant à la somme des trois nombres écrits sur les illustrations) peuvent être interprétées de cette façon. En revanche, les autres réponses, comme celles par exemple supérieures à 8 et 47, sont difficilement interprétables sans autre information.

Résultats des élèves

À partir de ces éléments d'interprétation, nous observons (*cf.* tableau 3) qu'une proportion importante des réponses peut être interprétée dans le cas des « petits nombres » (93 %) alors qu'un tiers des réponses (33 %) ne peut l'être dans le cas du problème avec des « grands nombres » (dans ce cas, nous précisons tout de même que 11 % des réponses (différentes de 47 et 84) sont supérieures à 37).

Transformation négative avec recherche de l'état final - « petits nombres ».		Transformation négative avec recherche de l'état final - « grands nombres ».	
Réponse correcte (4)	60 %	Réponse correcte (27)	41 %
Réponse correcte à 1 près (3 ou 5)	9 %	Réponse correcte à 1 près (26 ou 28)	10 %
Réponse 8	19 %	Réponse 47 *Réponse 46 ou 48 (pour information)*	14 % *2 %*
Réponse 14	5 %	Réponse 84	0 %
Absence de réponse	1 %	Absence de réponse	7 %
Autres réponses	6 %	Autres réponses	26 %

Tableau 3 *: Répartition des réponses aux problèmes de la figure précédente.*

Une telle interprétation des réponses de chacun des 83 élèves peut ainsi être menée sur chaque problème. Dans le cas où l'interprétation de la réponse est possible, elle permet de renseigner sur la réussite ou sur une des potentielles sources d'erreur (sous-processus de modélisation ou/et sous-processus de production de résultat), et ainsi, pour l'enseignant, lui permettre de proposer une régulation adaptée aux besoins de l'élève.

Il peut ainsi être intéressant pour un même élève de mettre en relation ses réponses et de rechercher si des erreurs du même type se retrouvent d'une classe de problèmes à l'autre ; ce qui peut permettre de révéler des démarches superficielles lorsque, par exemple :

- le choix de l'opération se fait en fonction uniquement de mots inducteurs, c'est-à-dire une somme (ou une différence) lorsque le mot « verse » (ou « prends ») apparaît dans l'énoncé ; ce qui conduit à une réponse fausse dans le cas de recherche de la transformation ;
- l'élève ajoute les trois nombres présents dans l'énoncé et dans la bande dessinée.

Par ailleurs, des réponses récurrentes à plus ou moins 1 ou 2 près de la réponse exacte peuvent provenir d'erreurs de calcul. Les feedbacks que devra apporter l'enseignant ne seront donc pas de même nature dans ces différents cas.

4.3. Analyse par classe de problèmes selon PN et GN

Intéressons-nous désormais à la performance des élèves selon la classe de problèmes en croisant la réussite aux problèmes avec petits et grands nombres. Cette analyse devrait nous permettre de déterminer l'effet de la variable « taille des nombres » sur le processus de modélisation sans tenir compte des erreurs de calcul ou de dénombrement ; pour cette raison, nous avons considéré comme réussite non seulement la réponse exacte au problème, mais aussi les réponses données à plus ou moins un près. L'ensemble des résultats est produit en annexe 2 et nous illustrons ci-après l'intérêt de cette analyse à partir des deux problèmes de recherche de l'état final après transformation négative (énoncés fournis figure 3).

Problèmes de transformation, transformation négative, recherche de l'état final		
	Réussite grands nombres	**Échec grands nombres**
Réussite petits nombres	43,4 %	25,3 %
Échec petits nombres	7,2 %	24,1 %

Tableau 4 *: Répartition des élèves selon la réussite/échec aux problèmes de la figure 2.*

Pour les 24 % des élèves qui échouent au problème à la fois avec des petits et des grands nombres, il faudrait rechercher plus précisément les causes de ces erreurs (par exemple à partir des erreurs types définies dans l'analyse *a priori*), mais la seule détermination de leur échec aux deux problèmes, ne permet pas à l'enseignant d'avoir un levier pour les faire progresser. Signalons à nouveau, que pour cette classe de problèmes, un élève peut répondre correctement grâce à des démarches superficielles basées uniquement sur les mots inducteurs ou sur le sens de la flèche proposée dans l'énoncé et sans activité dans le processus de modélisation.

Dans ces analyses croisées, deux catégories nous intéressent particulièrement.

a) *Les élèves qui réussissent avec des petits nombres mais échouent avec les plus grands.*

   Pour les 25,3 % des élèves qui relèvent de ce profil, l'enseignant pourrait s'appuyer sur leur capacité à réussir les problèmes avec des petits nombres pour qu'ils progressent sur la résolution avec de grands nombres, à condition de s'assurer que les échecs ne proviennent pas d'une erreur de calcul et que la réussite avec les petits nombres ne provient pas d'une démarche superficielle ;

b) *Les élèves qui échouent avec les petits nombres mais réussissent avec les grands nombres (7,2 % des élèves, soit 6 élèves).*

   Parmi ces six élèves, quatre donnent pour réponse 8 pour le problème avec petits nombres, c'est-à-dire la somme des nombres écrits sur les boîtes dessinées. Pour ces 4 élèves, l'enseignant pourrait enquêter sur les raisons de démarches superficielles déclenchées par la présence de petits nombres, alors que les grands nombres n'ont pas eu cet effet.

L'analyse *a priori* des tâches, qui a guidé les choix des valeurs numériques (afin de limiter des erreurs dues au calcul) et qui a conduit à une certaine interprétation des réponses, ne laissait pas présager de tels « profils » d'élèves. Nous précisons que ces profils apparaissent pour chacune des autres classes de problèmes (*cf.* annexe 2) et qu'ils représentent jusqu'à 16,9 % des élèves. Il y a évidemment des limites à la recherche de certaines cohérences dans le fonctionnement cognitif des élèves avec seulement deux problèmes, mais cela soulève différents types de questions : ces élèves se sont-ils redéfinis la tâche attendue (Rogalski, 2003), ce qui questionne les modalités de passation ? Les connaissances d'élèves de CP sont-elles suffisamment stables pour que de tels croisements de données soient pertinents ? Certains élèves ne déploient-ils pas des démarches superficielles différentes selon la taille des nombres ?

Conclusion

Nous rappelons que le test conçu dans le cadre de l'ingénierie évaluative et ses résultats (quantitatifs et qualitatifs) ont pour objectif permettre d'outiller l'enseignant pour l'évaluation des connaissances de ses élèves.

Pour cette raison, nous avons choisi de passer une évaluation dans les classes, de façon collective et avec des modalités que nous avons voulues reproductibles dans les conditions d'exercices « ordinaires », à la différence de nombreuses recherches qui ont été menées de façon clinique sur la résolution de problèmes avec des élèves de cet âge. En procédant ainsi, et malgré la vigilance didactique dont nous avons fait preuve pour concevoir et mettre en œuvre une épreuve valide, nous avons constaté la présence de résultats étonnants du point de vue de notre analyse didactique *a priori*, en particulier lorsque nous avons cherché à croiser les réussites/échecs selon la présence de petits ou grands nombres sur une même classe de problèmes.

Concernant la forme de l'énoncé du problème (texte + illustration), nous avons observé que le fait de présenter une séquence d'images chronologiques de ce qui a été lu et montré (*cf.* figure 2) a donné l'occasion à certains élèves de se redéfinir une tâche qui n'est pas celle attendue (faire la somme des trois nombres présents dans les illustrations). D'autres facteurs exercent une influence sur la compréhension de la tâche par les élèves. Certains sont communs à toutes les classes de CP, comme les capacités d'attention que l'on peut attendre des élèves de cet âge ou encore leurs connaissances liées aux premières fréquentations des problèmes, des nombres, du calcul et de l'écrit. D'autres peuvent avoir trait à certaines spécificités des classes testées, ici des élèves de REP. Le contrat didactique lors de la résolution de problèmes (Mounier, 2016) joue aussi sur la redéfinition que certains élèves se font de la tâche « résoudre un problème », qui serait par exemple de faire un calcul (mental/posé) et donner une réponse numérique sans s'interroger sur le choix d'un modèle mathématique en relation avec l'énoncé du problème, sans interpréter le résultat avant de fournir une réponse et/ou sans exercer de contrôle durant le processus de résolution. Une telle redéfinition de la tâche, qui conduit à des démarches superficielles, est particulièrement importante à déceler pour l'enseignant, d'autant plus que ce type de démarches ne conduit pas nécessairement à une réponse erronée.

Nos résultats, tant qualitatifs que quantitatifs, permettent, en appui avec le modèle adapté de Verschaffel et de Corte (*cf.* figure 1), de revenir sur certaines hiérarchies de réussite et de sensibiliser les enseignants sur deux points : l'insuffisance de la seule trace écrite (quand elle est présente) et la limite de l'interprétation de la réponse de l'élève en termes de réussite/échec pour évaluer ses connaissances et compétences en résolution de problèmes.

C'est alors en observant l'élève et en s'entretenant avec lui que l'enseignant pourrait ensuite davantage comprendre sa démarche et, par la suite, l'aider à progresser en lui fournissant les rétroactions adaptées. Il s'agit ainsi, pour l'enseignant, d'ajuster ses gestes évaluatifs (Blanchouin *et al.*, 2022b), qu'ils soient de recueil d'informations et/ou de rétroaction lors de la résolution de problèmes.

Si l'analyse *a priori* et le croisement des résultats sur une même classe de problèmes avec petits et grands nombres permettent d'apporter un premier éclairage sur la réussite des élèves et envisager des « profils », il reste cependant difficile de repérer les démarches superficielles, notamment lorsqu'elles conduisent à la bonne réponse. Selon nous, à la lumière des résultats obtenus ici et de nos lectures sur la résolution de problèmes, une piste pour l'enseignant pourrait être de proposer des énoncés de problèmes avec des analogies de substitution non-facilitatrices (Sander, 2018), par exemple un énoncé avec la présence du verbe « perd » mais pour lequel il y a nécessité de calculer une addition. Une autre piste pour l'enseignant serait d'amener l'élève à adopter une posture de « chercheur » (notamment en effectuant des contrôles de son activité), du moins à ne pas se redéfinir la tâche « résoudre un problème » comme « combiner les nombres en présence à l'aide d'une opération » ; différentes pratiques, comme l'utilisation d'un cahier de mathématicien.n.e ou une pratique régulière de problèmes « pour apprendre à chercher » peut permettre une évolution (Blanchouin *et al.*, 2023). Certains élèves ont pu cependant déployer une activité participant au sous-processus de modélisation, activité qui n'a pas abouti à un bon modèle, notamment du fait de la présence de « grands » nombres. Rivier, Scheibling-Seve et Sander (2022, p. 106) évoquent alors l'intérêt de considérer les analogies de simulation *« lorsque les valeurs numériques de l'énoncé ne permettent pas à la simulation mentale de conduire au résultat »* et indiquent les *« effets bénéfiques d'une intervention destinée à favoriser cet usage des principes arithmétiques pour résoudre des problèmes de soustraction et d'additions »*.

Enfin, en tant que chercheurs et pour comprendre plus précisément les démarches de l'élève, nous avons reconduit une telle expérimentation, en proposant un ensemble de problèmes semblable à celui présenté dans cet article, mais en passation individuelle, en présence du chercheur (Mounier *et al.*, à paraître). Dans ce nouveau protocole, les élèves ont été filmés en train de résoudre le problème et nous nous sommes entretenus avec eux à la suite de la résolution. L'analyse croisée des données issues des vidéos et des entretiens pour un même élève sur l'ensemble des problèmes a fait l'objet d'une méthodologie spécifique qui semble prometteuse puisque, sur une étude de cas, nous avons pu montrer le rôle déterminant que joue la maîtrise des procédures de calcul lors du sous-processus de modélisation. Par exemple, un élève réussissait presque tous les problèmes avec les petits nombres : il effectuait ses calculs et dénombrements avec les doigts. Pour les mêmes problèmes avec les grands nombres, il semblait emprunter une démarche superficielle additionnant systématiquement et immédiatement les deux nombres du problème. Nous la mettons en relation avec le fait qu'il ne maîtrisait une technique opératoire que pour l'addition (et pas pour les autres opérations).

Pour conclure, l'étude de la validité d'un test en résolution de problèmes est particulièrement complexe à ce stade de la scolarité, puisque, au-delà du choix des classes de problèmes et des valeurs de variables didactiques (contexte, valeurs numériques, etc.), de la rédaction de l'énoncé et des modalités de passation, se pose la question clé de l'interprétation de la réponse de l'élève au regard du modèle choisi. Si l'étude que nous avons menée ne permet pas toujours d'apporter des réponses sur l'interprétation des données recueillies, elle propose néanmoins des pistes de réflexion à exploiter pour l'enseignant et pour les chercheurs et elle interroge indirectement la validité de nombreux dispositifs d'évaluation existants. Les pistes et la méthodologie que nous proposons pourraient être exploitées y compris dans des évaluations nationales et internationales pour pouvoir interroger finement la singularité des connaissances et démarches d'un élève.

Références bibliographiques

Blanchouin, A., Grapin, N., Mounier, É., Marques, C., Marques, S. & Prigent, L. (2023). Étude du travail collectif entre professeurs des écoles et chercheurs/formateurs autour de trois dispositifs d'enseignement concernant la résolution de problèmes au cycle 2. *Troisième Congrès international de la Théorie de l'Action Conjointe en Didactique (TACD)*. 7,8,9 novembre à Brest.

Blanchouin, A., Grapin, N., Mounier, É. & Sayac, N. (2022a). Les pratiques d'évaluation en mathématiques à l'école élémentaire : deux dispositifs de recherche formation. *Éducation & didactique. Numéro spécial LEA, 15-3*, 65-84.

Blanchouin, A., Grapin, N. & Mounier, É. (2022b). Documenter l'activité évaluative des professeurs des écoles à partir de leurs gestes évaluatifs - étude de cas en mathématiques. Évaluer. *Journal international de recherche en éducation et formation, 8*(1), 3-28. https://doi.org/10.48782/e-jiref-8-1-3

Butlen, D. (2007). *Le calcul mental entre sens et technique*. Besançon : Presses universitaires de Franche-Comté.

Chenevotot, F. Ledan, L., Blanchouin, A., Beylot, D., Mounier, É. & Grapin, N. (2022). Analyser des connaissances en calcul mental additif au début du cycle 2 pour mieux comprendre l'activité de l'élève en résolution de problèmes. *Actes du 48^e^ colloque Copirelem* (pp. 234-257). Toulouse, 14-16 juin 2022.

Conne, F. (1989). Comptage et écritures en ligne d'égalités numériques. *Recherches en didactique des mathématiques, 9*(1), 71-115.

Fagnant, A. (2005). Résoudre et symboliser des problèmes additifs et soustractifs en début d'enseignement primaire. Dans M. Crahay, L. Verschaffel, E. de Corte & J. Grégoire (éds.), *Enseignement et apprentissage des mathématiques : que disent les recherches psychopédagogiques ?* (pp. 131-150). Bruxelles : De Boeck.

Fagnant, A. (2013). Opérations arithmétiques et symbolisations variées. Partir des démarches informelles des élèves pour donner du sens aux apprentissages. *Éducation & formation, e-298-01*, 23-38.

Fagnant, A. (2018). Des illustrations qui accompagnent les problèmes à la construction de représentations schématiques par les élèves : quels enjeux face aux problèmes standards et problématiques ? Dans J. Pilet & C. Vendeira (dir.), *Actes du séminaire de didactique des mathématiques de l'ARDM* (pp. 94-113). Paris : IREM de Paris.

Fayol, M. (2008). La résolution de problèmes : de la compréhension aux opérations. *Actes du séminaire national - L'enseignement des mathématiques à l'école primaire* (pp. 49-60). 13-14 novembre 2007. MEN-DGESCO.

Feyfant, A. (2015). La résolution de problèmes de mathématiques au primaire. *Dossier de veille de l'IFÉ, 105*, novembre. Lyon : ENS de Lyon.

Grapin, N., Chenevotot, F., Ledan, L., Beylot, D., Mounier, É. & Blanchouin, A. (2022). Étude exploratoire de procédures d'élèves de 7-8 ans en calcul mental aditif. *Revue Math-École, 238*, 29-40.

Grapin N. & Grugeon-Allys, B. (2018). Approches psychométrique et didactique de la validité d'une évaluation externe en mathématiques : quelles complémentarités ? Quelles divergences ? *Mesure et évaluation en éducation, 41-2*, 37-66.

Grapin, N. & Mounier, É. (2019). Concevoir et mettre en œuvre des évaluations au service des apprentissages numériques des élèves au cycle 2. Dans S. Coppé, E. Roditi (éds.), *Actes de l'école d'été de didactique des mathématiques, vol. 2* (pp. 391-410). Paris, août 2017.

Grapin N., Vantourout, M. & Grugeon-Allys, B. (2022). S'assurer de la validité des évaluations en milieu scolaire. *Recherches en didactiques, 33*, 59-83.

Houdement, C. (2017). Résolution de problèmes arithmétiques à l'école. *Grand N, 100*, 59-78.

Mazollier, M.-S., Mounier, É. & Pfaff, N. (2016). *Mon année de Maths CP*. Fichier élève et manuel pédagogique. Paris : Editions SED.

Mottier-Lopez, L. (2015). *Évaluations formative et certificative des apprentissages*. Bruxelles : de Boeck.

Mounier, É., Beylot, D., Blanchouin, A., Chenevotot, F., Grapin, N. & Ledan, L. (à paraître). Repérer les démarches en résolution de problèmes d'élèves de grade 2 par l'analyse de leurs procédures : méthodologie et étude de cas. *Annales de didactique et de sciences cognitives*.

Mounier, É. (2016). Évaluer les connaissances mathématiques dans des tâches de dénombrement chez les élèves du Cours Préparatoire, élèves âgés de 6-7 ans : un nouvel outil d'analyse, des premiers résultats. *Actes du colloque de l'Association pour le Développement des Méthodologies d'Evaluation en Education ADMEE Europe*. Lisbonne : ADMEE.

Mounier, É. (2010). *Numération au CP. Vers de nouvelles pistes*. [Thèse de doctorat, Paris 7].

Riley, M., Greeno, J. & Heller, J. (1984). Development of children's problem-solving ability in arithmetic. Dans H. Ginsburg (dir.), *The development of mathematical thinking* (pp. 153-196). New York : Academic Press.

Rivier, C., Scheibling-Seve, C. & Sander, E. (2022). Études des types de problèmes arithmétiques à énoncés verbaux proposés dans 12 manuels scolaires français de cycle 2 : concordance et discordance par rapport à trois formes d'analogies. *Revue française de pédagogie, 216*, 101-116.

Rogalski, J. (2003). Y a-t-il un pilote dans la classe ? Une analyse de l'activité de l'enseignant comme gestion d'un environnement dynamique ouvert. *Recherches en Didactique des Mathématiques, 23*(3), 343-388.

Sander, E. (2018). Une perspective interprétative sur la résolution de problèmes arithmétiques : le cadre A-S3. Dans J. Pilet & C. Vendeira (dir.), *Actes du séminaire de didactique des mathématiques de l'ARDM* (pp. 94-113). Paris : IREM de Paris.

Sayac, N. (2019). Approche didactique de l'évaluation et de ses pratiques en mathématiques. *Recherches en didactique des mathématiques, 39*(3), 283-329

Van Dooren, W., Verschaffel, L., Greer, B., De Bock, D. & Crahay, M. (2015). La modélisation des problèmes mathématiques. *Psychologie des apprentissages scolaires, 8*, 199-219.

Vantourout, M. (2020). Évaluations et didactiques : quelles synergies ? Dans N. Younes, C. Gremion & E. Sylvestre (dir), *Évaluations, sources de synergies* (pp. 99-127). Neuchâtel (Suisse) : Presses de l'ADMEE.

Vantourout, M. & Maury, S. (2017). Évaluation de la lecture au CP : mise en oeuvre d'une approche multidimensionnelle. *Éducation et didactique, 11*(1), 45-62.

Verschaffel, L. & De Corte, E. (1997). Word problems : A vehicle for promoting authentic mathematical understanding and problem solving in the primary school? Dans T. Nunes & P. Bryant (dir.), *Learning and Teaching Mathematics: an International Perspective* (pp. 69-97). UK : Psychology Press Ltd.

Verschaffel, L. & De Corte, E. (2008). La modélisation et la résolution des problèmes d'application : de l'analyse à l'utilisation efficace. Dans M. Crahay, L. Verschaffel, E. De Corte & J. Grégoire (dir.), *Enseignement et apprentissage des mathématiques. Que disent les recherches psychopédagogiques ?* Bruxelles : De Boeck.

Vergnaud, G. (1991). La théorie des champs conceptuels. *Recherches en didactique des mathématiques, 10*(2/3), 133-169.

DEPP (2021). Premiers résultats de l'enquête sur les pratiques d'enseignement des mathématiques, Praesco, en classe de CM2 en 2019. *Note d'information 21.10*.
https://www.education.gouv.fr/premiers-resultats-de-l-enquete-sur-les-pratiques-d-enseignement-des-mathematiques-praesco-en-classe-309564

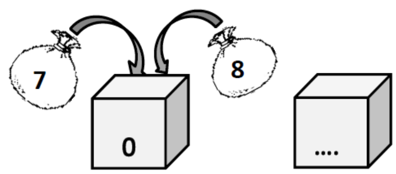
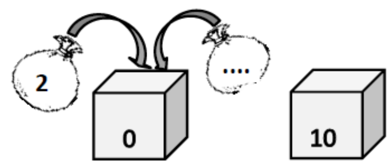
DEPP (2023). Évaluations 2023 - Point d'étape CP. Premiers résultats. *Série Études Document de travail*, 2023-E01.

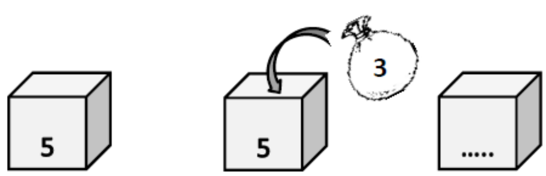
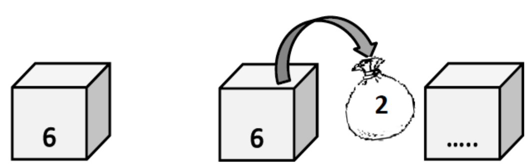
MENJS (2016). *Principes d'action pour évaluer les acquis des élèves*.
https://eduscol.education.fr/141/modalites-d-evaluation-des-acquis-scolaires-des-eleves

MENJS (2019). *Programme d'enseignement du cycle des apprentissages fondamentaux (cycle 2)*.

MENJS (2018). *Pour enseigner les nombres, le calcul et la résolution de problèmes au CP*.

Annexe 1
Ensemble des problèmes posés

Problèmes de composition de mesure (réunion ; partie/tout)	
Recherche du tout	**Recherche d’une partie**
Petits nombres : *On cherche le nombre de jetons dans la boîte à la fin.* J’ai deux sachets et une boite vide. Dans un sachet il y a 7 jetons d’une couleur. Dans l’autre sachet il y a 8 jetons d’une autre couleur. Je verse les deux sachets dans la boîte vide. **Quel est le nombre de jetons dans la boîte à la fin ?** *Écris la réponse sur les pointillés.* 	Petits nombres : *On cherche le nombre de jetons dans le sachet mystère.* J’ai deux sachets et une boite vide. Dans un sachet il y a 2 jetons d’une couleur. Dans l’autre sachet il y a des jetons d’une autre couleur. Je ne sais pas combien il y en a. C’est le sachet mystère. Je verse les deux sachets dans la boîte vide. A la fin, j’ai 10 jetons dans la boîte. **Quel est le nombre de jetons qu’il y avait dans le sachet mystère ?** *Écris la réponse sur les pointillés.* 
Grands nombres : même énoncé avec 35 jetons d’une couleur et 10 de l’autre.	Grands nombres : même énoncé avec 30 jetons d’une couleur et 52 jetons en tout (dans la boîte).

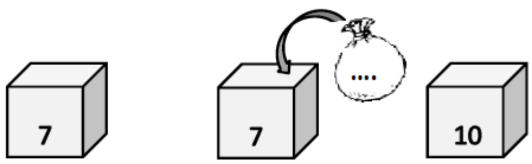
Problèmes de transformation d’état, recherche de l’état final	
Transformation positive	**Transformation négative**
Petits nombres : *On cherche le nombre de jetons dans la boîte à la fin.* Au début, il y a 5 jetons dans la boîte. Je mets encore 3 jetons dans la boite. **Quel est le nombre de jetons dans la boîte à la fin ?** *Écris ta réponse sur les pointillés.* 	Petits nombres : *On cherche le nombre de jetons dans la boîte à la fin.* Au début, il y a 6 jetons dans la boîte. Je prends 2 jetons dans la boite et je les mets dans un sachet. **Quel est le nombre de jetons dans la boîte à la fin ?** *Écris ta réponse sur les pointillés.* 
Grands nombres : même énoncé avec 13 jetons dans la boîte et on en met encore 16.	Grands nombres : même énoncé avec 37 jetons dans la boîte et on en prend 10.

<table>
<tr><th colspan="2">Problèmes de transformation d’état, recherche de l’état final</th></tr>
<tr><th>Transformation positive</th><th>Transformation négative</th></tr>
<tr><td>Petits nombres :

On cherche le nombre de jetons que l’on verse dans la boîte.

Au début, il y a 7 jetons dans la boîte.
Je verse tous les jetons du sachet dans la boîte.
A la fin, dans la boîte, il y a 10 jetons.

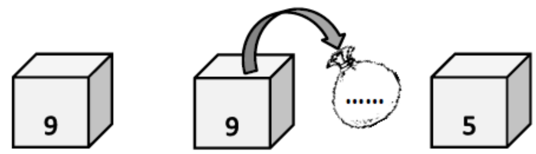
Quel est le nombre de jetons que j’ai versé dans la boîte ?
Écris ta réponse sur les pointillés.


</td>
<td>Petits nombres :

On cherche le nombre de jetons que l’on met dans le sachet.

Au début, il y a 9 jetons dans la boîte.
Je prends des jetons dans la boite et je les mets dans le sachet.
A la fin, dans la boîte, il y a 5 jetons.

Quel est le nombre de jetons que j’ai mis dans le sachet ?
Écris ta réponse sur les pointillés.


</td></tr>
<tr><td>Grands nombres : même énoncé avec 23 jetons dans la boîte au début et on en a 53 à la fin.</td><td>Grands nombres : même énoncé avec 46 jetons dans la boîte au début et on en a 46 à la fin.</td></tr>
</table>

Annexe 2
Croisement champ numérique - réussite

Problèmes de composition de mesure - recherche du tout		
	Réussite GN	**Échec GN**
Réussite PN	55 (66,3 %)	12 (14,5 %)
Échec PN	8 (9,6 %)	8 (9,6 %)

Problèmes de composition de mesure - recherche d’une partie		
	Réussite GN	**Échec GN**
Réussite PN	34 (41,0 %)	24 (28,9 %)
Échec PN	4 (4,8 %)	21 (25,3 %)

Problèmes de transformation, transformation positive, recherche de l’état final		
	Réussite GN	**Échec GN**
Réussite PN	33 (39,8 %)	22 (26,5 %)
Échec PN	14 (16,9 %)	14 (16,9 %)

Problèmes de transformation, transformation négative, recherche de l’état final		
	Réussite GN	**Échec GN**
Réussite PN	36 (43,4 %)	21 (25,3 %)
Échec PN	6 (7,2 %)	20 (24,1 %)

Problèmes de transformation, transformation positive, recherche de la transformation		
	Réussite GN	**Échec GN**
Réussite PN	24 (28,9 %)	35 (42,2 %)
Échec PN	2 (2,4 %)	22 (26,5 %)

Problèmes de transformation, transformation négative, recherche de la transformation		
	Réussite GN	**Échec GN**
Réussite PN	26 (31,3 %)	32 (38,6 %)
Échec PN	4 (4,8 %)	21 (25,3 %)