IMPACT DES ILLUSTRATIONS ACCOMPAGNANT LES PROBLEMES PROBLEMATIQUES : FOCUS SUR LE CODAGE DES REPONSES

FAGNANT[*] Annick –AUQUIERE[**] Amélie

Résumé – En prolongement des travaux menés par Dewolf et al. (2017), la présente étude vise à évaluer l'impact de trois types d'illustrations sur la proportion de réponses réalistes produites par des élèves de 10-12 ans face à des problèmes problématiques. Cet article se centre sur une analyse exploratoire portant sur l'établissement du codage des réponses en « réalistes » *vs* « non réalistes ». Les productions d'élèves mettent en lumière la diversité des réponses produites et la difficulté de ce codage.

Mots-clefs : Evaluation, Problème problématique, Illustration, Réponse réaliste, Réponse non réaliste

Abstract – Following the work conducted by Dewolf et al. (2017), the present study intends to assess the impact of three types of illustrations on the proportion of realistic reactions produced by 10-12 years' old pupils confronted with problematic problems. This paper focuses on an exploratory analysis aiming at establishing the coding between "realistic" or "non realistic" reactions. Pupils productions highlight the diversity of the answers and the difficulty of this coding.

Keywords: Assessment, Problematic problem, Illustration, Realistic Reaction, Non-realistic Reaction

I. INTRODUCTION

En France, au début des années 1980, plusieurs études ont montré que la plupart des élèves fournissent une réponse numérique précise face à des problèmes pourtant insensés, comme le célèbre problème de « l'âge du capitaine » : « *Il y a 26 chèvres et 10 moutons sur le bateau. Quel est l'âge du capitaine ?* » (voir Baruk, 1985). Un autre exemple connu provient d'une étude américaine dans laquelle le problème suivant est soumis à des élèves de 13 ans : « *Un bus de l'armée peut contenir 36 soldats. Si 1128 soldats doivent être conduits en bus à leur site d'entraînement, combien faudra-t-il de bus ?* ». Seuls 23% des élèves fournissent la réponse attendue (32 bus). Les autres n'interprètent pas correctement le résultat et proposent « 31 reste 12 », « 31,33 » ou « 31 » (Verschaffel, Greer & De Corte, 2000). Ces résultats étonnants peuvent sans doute en partie s'expliquer par une rupture du « contrat didactique » (Brousseau, 1990) dans la mesure où les problèmes proposés vont à l'encontre des « normes socio-mathématiques » (Yackel & Cobb, 1996) établies dans les classes, mais ils questionnent néanmoins sur ce qui peut conduire les élèves à procéder de cette façon. Ils questionnent aussi quant aux pratiques de classe qui conduisent les élèves à penser que le « contrat » est de répondre coûte que coûte à tout problème proposé en classe (comme dans le 1er exemple) ou à négliger d'interpréter la solution obtenue de façon réaliste (comme dans le 2e exemple).

Ces questionnements ont conduit les chercheurs à étudier de manière assez systématique ce phénomène de « perte de sens » (Van Dooren, Verschaffel, Greer, De Bock, & Crahay, 2015 ; Verschaffel & De Corte, 2008), ainsi que les « présupposés » qui les sous-tendent et les éléments pouvant être à l'origine de ces présupposés (Depaepe, De Corte, & Verschaffel, 2015). De nombreuses études ont aussi tenté de voir dans quelle mesure les élèves utilisaient leurs connaissances de la vie réelle pour résoudre ce type de problèmes (que les auteurs qualifient de « problématiques ») et quel est l'impact de certaines consignes ou conditions de passation de l'évaluation sur la propension des élèves à produire des réactions réalistes (Dewolf, Van Dooren, Ev Cimen, & Verschaffel, 2014 ; Dewolf, Van Dooren & Verschaffel, 2017 ; Mellone, Verschaffel, & Van Dooren, 2017). Parmi ces divers éléments, c'est à l'impact des illustrations accompagnant les problèmes que notre étude s'intéresse.

[*] Université de Liège – Belgique – afagnant@uliege.be
[**] Université de Liège – Belgique – amelie.auquiere@uliege.be

Dans le cadre de cet article, après avoir présenté le cadrage théorique permettant de cibler la problématique et avoir expliqué le dispositif de recherche établi, c'est à une analyse exploratoire portant sur l'établissement du codage en réponses « réalistes » *vs* « non réalistes » que nous allons nous intéresser. Un focus sur cette étape nous a en effet semblé constituer une porte d'entrée intéressante pour ce *SPE* centré sur l'évaluation.

II. CADRAGE THEORIQUE ET PROBLEMATIQUE DE RECHERCHE

Les deux études pionnières dans ce domaine ont été réalisées dans les années 1990 en Irlande du Nord et en Communauté flamande de Belgique. Elles ont été reproduites dans de nombreux pays et aboutissent toujours à des résultats assez similaires (voir Verschaffel et al., 2000 pour une synthèse). Dans l'étude de Verschaffel, De Corte et Lasure (1994), menée en 5^{e} année primaire (*grade 5*), un test papier-crayon composé de 10 paires de problèmes a été proposé aux élèves. Chaque paire d'items est composée d'un « problème standard », qui peut être résolu par l'application directe d'une opération, et d'un « problème problématique » pour lequel la simple application d'une opération pose question à partir du moment où des connaissances réalistes liées à la situation sont prises en considération.

Le tableau 1 présente trois paires de problèmes et des exemples de réponses réalistes (assorties de la proportion observée) pour les problèmes problématiques.

Problèmes standards	**Problèmes problématiques** RR = réponses faisant appel à des considérations réalistes	
Steve a acheté 5 planches de 2 m chacune. Combien de planches de 1 m peut-il faire à partir de ces planches ?	Steve a acheté 4 planches de 2,5 m chacune. Combien de planches de 1 m peut-il faire à partir de ces planches ?	**RR : 14%** - *Ex. Steve peut faire 8 planches (ou 10, mais avec de la bonne colle !).*
Un bateau navigue à une vitesse moyenne de 45 km/h. Combien de temps lui faudra-t-il pour parcourir 180 km ?	John court le 100 m en 15 secondes. Combien de temps lui faudra-t-il pour parcourir un km ?	**RR : 3%** - *Ex. Certainement plus de 150 secondes.*
Christophe fait une balade. Il marche 8 km durant la matinée et 15 km durant l'après-midi. Combien de km a-t-il marché en tout ?	Bruce et Alice vont à la même école. Bruce habite à 17 km de l'école et Alice à 8 km de l'école. Quelle est la distance entre la maison de Bruce et la maison d'Alice.	**RR : 5%** - *Ex. La réponse doit se situer entre 9 et 25 km.*

Tableau 1 *– Exemples de problèmes standards et problématiques, ainsi que de réponses réalistes associées*
(d'après l'étude de Verschaffel et al., 1994, voir aussi Verschaffel et al., 2000)

Les élèves produisent très peu de réponses réalistes. Ils ont tendance à appliquer « envers et contre tout » des opérations arithmétiques en vue de fournir des réponses numériques précises à tous les problèmes proposés, ce que les auteurs qualifient de « réactions non réalistes » (Verschaffel et al., 2000). L'extrait suivant explicite la distinction entre réponses « réalistes » (RR) *vs* « non réalistes » (NR) face aux problèmes problématiques (P-items).

> A reaction was scored as a NR when it was the result of a straightforward execution of the mathematical operation, without any further comment about the problematic nature of the problem or real-life consideration that might jeopardize the appropriateness of the executed operation(s). In contrast, when the answer—a precise numerical answer, an answer indicating some kind of estimation, or an answer stating that the problem is unsolvable—was the result of the use of real-world knowledge related to the realistic modelling issue involved in the P-item, it was considered as a RR. Also when a pupil gave a straightforward non-realistic answer but made an additional comment that contained a trace of a realistic consideration about the modelling issue involved in the P-item, it was considered as a RR (Dewolf et al., 2017, p. 337).

Plusieurs études ont tenté d'amener les élèves à construire un « modèle de situation » plus riche (c.-à-d. prenant en compte des connaissances de la vie réelle et visant à diminuer la tendance à répondre de façon non réaliste) en « manipulant » différentes conditions de passation. Ainsi, certaines ont introduit les problèmes en proposant un *avertissement* visant à attirer l'attention des élèves quant au fait que le test contenait certains problèmes un peu particuliers ; d'autres ont proposé une tâche préalable invitant les élèves à *reformuler* et à enrichir le problème (seul ou en duo) et d'autres encore ont *accompagné les problèmes d'illustrations* ayant pour objectif d'enrichir les éléments concrets pris en compte dans la construction du modèle de situation (Dewolf et al., 2014, 2017 ; Mellone et al., 2017).

Dans une première étude, Dewolf et al. (2014) ont proposé des problèmes problématiques à des élèves de 5e primaire (*grade 5*) selon quatre conditions de passation : problèmes accompagnés d'un *avertissement*, problèmes accompagnés d'une *illustration*, problèmes accompagnés de ces deux types d'indice et problèmes seuls. Contrairement à leurs attentes, aucun des deux types d'indice n'a d'effet, tout comme d'ailleurs la combinaison de ceux-ci.

Prolongeant ces travaux et émettant l'hypothèse que les illustrations utilisées n'étaient pas suffisamment significatives, Dewolf et al. (2017) ont alors comparé l'effet de trois types d'illustrations en soumettant 7 items problématiques à 288 élèves de 5e et 6e années primaires (*grades 5-6*). S'appuyant sur la typologie d'Elia et Philippou (2004), ils décrivent les illustrations retenues dans leur étude comme étant "*représentationnelles*" au sens où elles décrivent un contexte réaliste et représentent certains éléments de contenu évoqués dans le problème *(≠ décoratives*), qu'elles n'orientent pas la stratégie de résolution (≠ *organisationnelles*) et qu'elles ne contiennent pas de données essentielles pour résoudre le problème (≠ *informationnelles*). Ce qui distingue alors les trois modalités comparées, c'est que le premier type d'illustrations représente simplement le contexte dans lequel se situent les problèmes (comme dans l'étude de 2014) alors que les deux autres contiennent un élément supplémentaire, ciblé sur la particularité de la situation d'un point de vue réaliste. Dans la troisième modalité, cet élément est mis en évidence en couleur (voir figure 1 pour un exemple des trois types d'illustration). Afin d'éviter un effet-classe, les trois types d'illustrations sont distribués aléatoirement au sein de chaque classe et ce sont les résultats globaux (toutes classes confondues) qui sont comparés par les auteurs. Dans chacune des conditions, les élèves reçoivent aussi deux avertissements visant à attirer leur attention sur le fait que certains problèmes sont un peu particuliers et les invitant à regarder attentivement les illustrations.

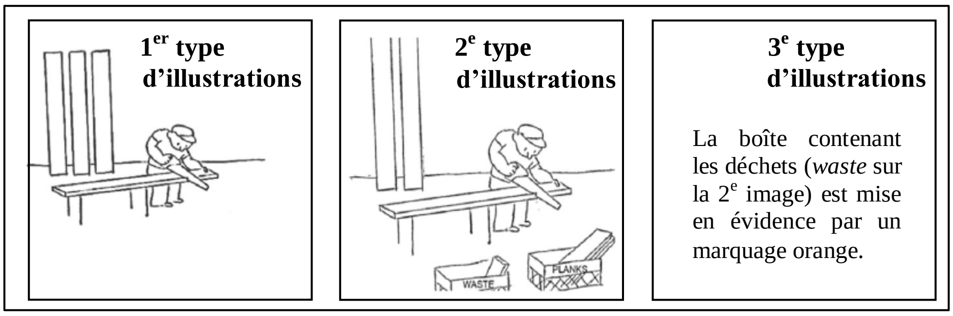

***Figure 1** – Les illustrations utilisées dans l'étude de Dewolf et al. (2017, p. 343) face au problème des planches dont l'énoncé est repris dans le tableau 1*

Les résultats obtenus par les chercheurs sont globalement décevants : les élèves produisent en moyenne seulement 23,7% de RR et il n'y a pas de différences significatives en fonction des trois conditions expérimentales. Les deux nouveaux types d'aide visuelle n'apportent donc aucune plus-value substantielle alors qu'ils visaient pourtant à mettre en exergue un élément-clé du problème, si l'on considère celui-ci d'un point de vue réaliste.

Parmi les hypothèses évoquées par les auteurs pour expliquer ces résultats, les présupposés (ou croyances) des élèves relatifs, non seulement à la résolution de problèmes en général, mais aussi au rôle et à l'importance des illustrations accompagnant les problèmes ont retenu notre attention. Ils estiment que les illustrations ont pu être sous-évaluées par les élèves qui ont pu les considérer comme purement décoratives et donc les ignorer ou les analyser très superficiellement. En effet, même si les auteurs les considèrent comme « *représentationnelles* » (au sens où elles évoquent un contexte réaliste), elles pourraient tout autant être considérées comme « *décoratives* » dans la typologie d'Elia et Philippou (2004) si elles accompagnaient un problème standard. En prolongement de l'étude de Dewolf et al. (2017), notre étude cherche à évaluer l'effet de la présence d'illustrations « *informationnelles* » qui, en présentant les données numériques de l'énoncé au sein même des illustrations (et uniquement dans celles-ci), vont contraindre les élèves à les prendre en compte pour résoudre le problème.

III. DISPOSITIF DE RECHERCHE ET QUESTIONNEMENT PREALABLE

Le dispositif expérimental mis en place est tout à fait parallèle à celui de l'étude précitée : nous avons utilisé les 7 mêmes problèmes problématiques ; le double *avertissement* a été donné à tous les élèves et trois types d'illustrations ont été répartis aléatoirement dans chaque classe. C'est au niveau des types d'illustrations utilisés que se situe la différence majeure : les deux premiers types correspondent à ceux utilisés par les chercheurs flamands (voir figure 1) et le troisième type transforme le type 2 en retirant les données du texte de l'énoncé pour les placer sur le dessin (voir figure 2).

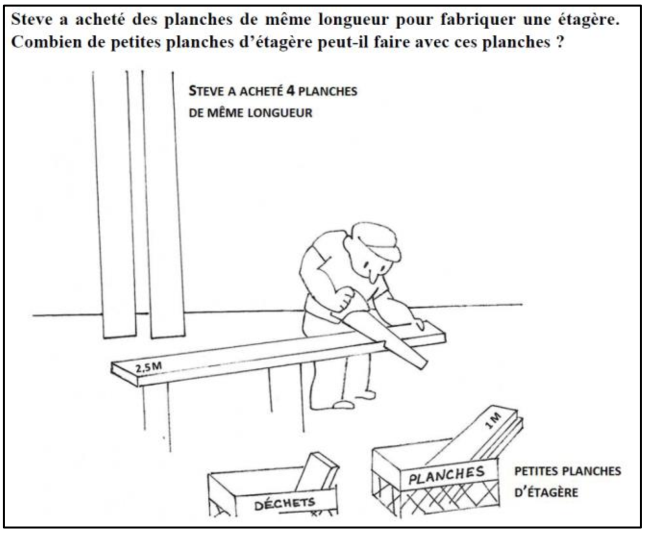

Figure 2 *– Exemple d'illustration informationnelle utilisée dans notre étude pour le problème des planches*

Dans ce cas, l'élève ne peut pas se contenter de lire le texte et de jeter (ou pas) un rapide coup d'œil sur l'illustration. Il est obligé de prendre en compte celle-ci pour y retirer les données utiles à la résolution du problème. L'hypothèse que nous formulons est alors que la contrainte de prise en compte de l'illustration impactera positivement la production de réponses réalistes. Autrement dit, nous faisons l'hypothèse que les *illustrations informationnelles* conduiront les élèves à plus de réalisme que les *illustrations représentationnelles*.

Pour tenter de vérifier cette hypothèse, le test a été soumis à un échantillon de près de 500 élèves provenant d'une vingtaine de classes de 5^{e}-6^{e} années primaires (*grades 5-6*) situées en

Belgique francophone. Les données ont été récoltées dans le cadre de travaux pratiques proposés à des étudiants inscrits au master en Sciences de l'Education dans notre institution.

Dans les différentes études centrées sur cette thématique, les auteurs catégorisent l'ensemble des réponses produites par les élèves en deux catégories (RR *vs* NR) auxquelles s'ajoutent les (rares) omissions. Le pourcentage de RR est alors calculé sur la somme des réponses produites. Par exemple, dans l'étude de Dewolf et al. (2017) dans laquelle 288 élèves ont répondu chacun à 7 problèmes *(ce qui conduit à 2016 réponses possibles)*, les auteurs précisent que "*The percentage of RRs was calculated on a total of 2008 trials (since eight items were not solved). Across all conditions, only 23.8 % of the reactions was scored as realistic*" (p. 342). Bien que les explications et les exemples fournis dans les articles semblent clairs, la diversité des productions produites par les élèves face à ces problèmes soulève un certain nombre de questionnements quant au codage dichotomique en RR *vs* NR. Dans le cadre de cet article, c'est sur cette étape préalable à la vérification de l'hypothèse susmentionnée que nous allons nous focaliser, en proposant une analyse exploratoire centrée sur le problème des planches.

IV. QUELQUES EXEMPLES ISSUS D'UNE ANALYSE EXPLORATOIRE

En vue de réaliser une première analyse exploratoire, les étudiants qui ont recueilli les données ont été invités à les coder en 6 catégories : (1) réponse réaliste attendue (RR) ; (2) réponse non réaliste attendue (NR) ; (3) réponse non réaliste attendue, mais avec erreur technique (ET) ; (4) hésitation quant au caractère réaliste (RR/NR ?) ; (5) autres réponses (AR) et (6) omission (OM). C'est parce que nous avions anticipé quelques difficultés de codage que nous avions prévu ces catégories « RR/NR ? » et « AR ». En demandant aux codeurs de ne coder comme « RR » ou « NR » que les éléments correspondant « au plus près » aux réponses que nous avions pu anticiper grâce aux écrits antérieurs, nous cherchions à limiter les erreurs de codage. Nous avions aussi prévu un fichier partagé où les étudiants pouvaient introduire une réponse d'élève pour demander un avis sur son codage et ce, dans le but d'uniformiser le codage de certaines réponses posant question. Procéder avec ces codes multiples devait aussi permettre d'avoir une vue d'ensemble des données, avant de procéder nous-mêmes à un recodage « calibré » mettant en place diverses techniques en vue d'assurer sa fiabilité.

En codant ainsi tous les problèmes, nous nous attendions à obtenir une majorité de codes « RR » et « NR », traduisant les deux types de réponses attendus (les « ET » traduisant également des « NR », que nous pourrons d'emblée regrouper avec les autres). Les premières données « brutes » montrent que la réalité est toute autre : elles mettent en exergue la diversité des productions d'élèves et la difficulté de coder leurs réponses de façon dichotomique.

Dans le **problème des planches**, le premier codage brut des données conduit à identifier 14,7% de « RR », 36,5% de « NR » (incluant quelques « ET ») ; 0,6% de « RR/NR ? » ; 41% d'autres réponses et 7,3% d'omission. Que cache chacun de ces codes ?

Au niveau des **codes « RR »**, on rencontre des réponses « 8 planches » (avec ou sans indication des unités et avec ou sans trace du calcul effectué) et des réponses évoquant explicitement le caractère problématique lié à la construction de 10 planches (par ex. « *Avec toutes les planches ça fait 8 planches de 1m et il reste 4 morceaux de bois de 0,5 ; si on les recolle ça fait 2 planches de plus = 10 planches* »).

Les **codes « NR »** traduisent des réactions consistant à proposer comme réponse « 10 planches » (avec ou sans unité et/ou trace du calcul), sans que cette réponse soit accompagnée de la moindre indication de la difficulté que pourrait rencontrer le menuisier avec ses

« restes » d'un demi-mètre. Dans ces deux cas, il n'y a pas d'ambiguïté possible, si l'on se réfère aux explications données par les auteurs. Par contre, la réponse « *Pas clair parce que 4 planches de 2,5 m ok mais une à 1 m c'est pas possible* » est plus complexe à interpréter dans la mesure où l'on ne sait pas si l'élève cible bien où se situe la difficulté, ni même s'il a compris qu'on pouvait sans problème couper deux planches de 1 m au départ d'une planche de 2,5 m. Fait-il preuve d'un certain réalisme ou fait-il (simplement) preuve d'une incompréhension du problème ? Pour cette réponse, qui avait été partagée dans le fichier commun, le code provisoirement proposé est le **code « RR/NR ? »** indiquant la difficulté du positionnement dichotomique.

Les **codes « AR »** recouvrent une grande diversité. De façon clairement non-représentative (et non-exhaustive) et avec pour seule intention (à ce stade) d'illustrer la diversité des productions d'élèves codées « AR », la figure 3 en propose quatre exemples.

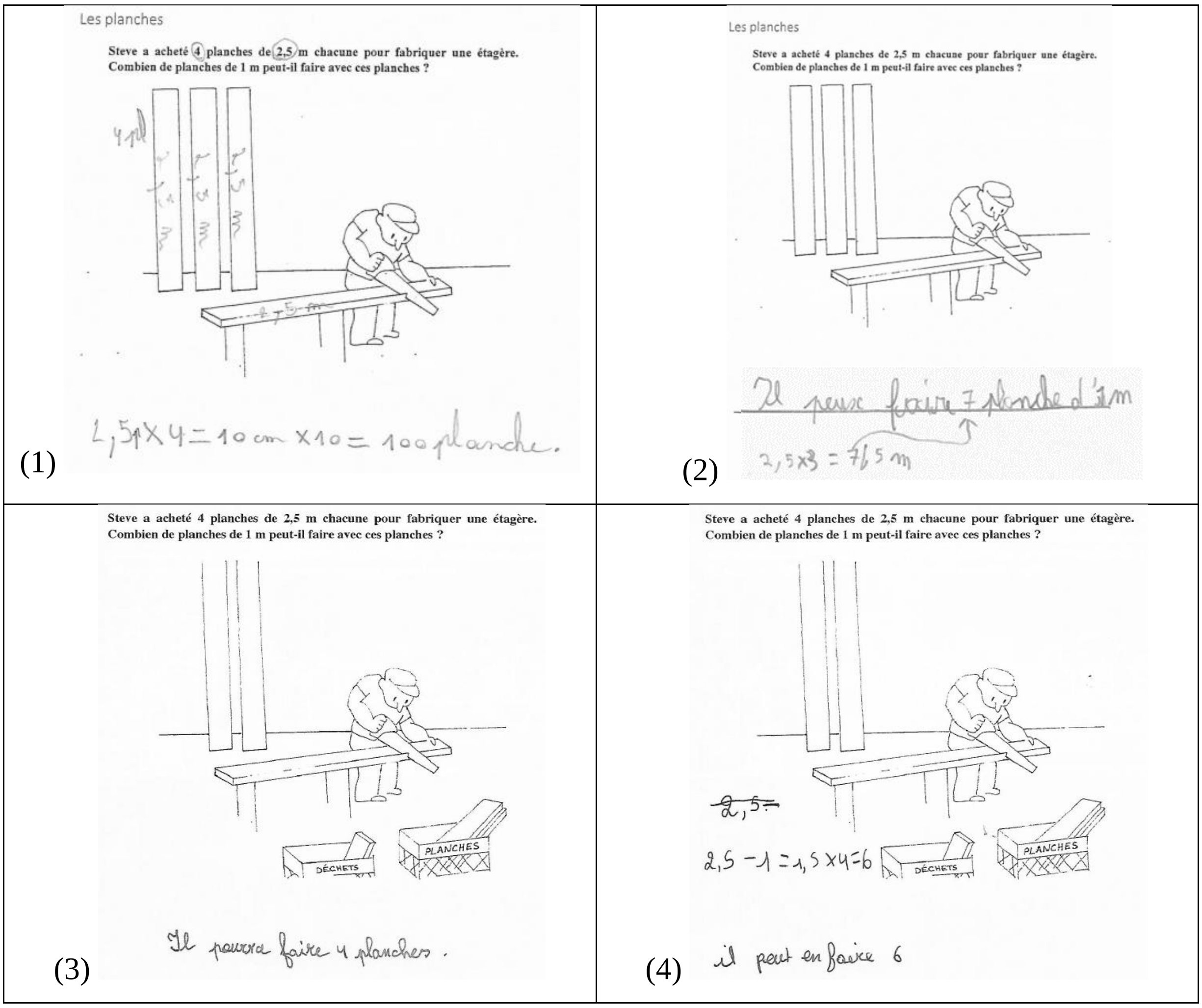

Figure 3 *– Quatre exemples de réponses codées « AR » (autres réponses) face au problème des planches*

Dans le premier cas (1), l'élève met en œuvre la démarche non réaliste attendue, mais il poursuit ensuite les calculs, sans doute parce qu'il s'embrouille dans les unités entre les centimètres et les mètres. *A priori*, ce type de réponse pourra être recodé comme « NR », dans la même logique que les réponses de type « 10 planches ». Dans le deuxième cas (2), l'élève a sans doute pris en compte l'illustration *représentationnelle* pour considérer uniquement les 3 planches posées contre le mur plutôt que les 4 planches indiquées dans l'énoncé. Son calcul « 2,5x3 » s'apparente à la démarche non réaliste attendue, mais le conduit à une réponse décimale (7,5 m et non 10 m). On pourrait penser qu'il répond ensuite de façon réaliste puisqu'il répond « 7 planches » (et non 7,5 planches), mais cette réponse nécessiterait tout de même de coller ensemble deux planches de 0,5 mètre, difficulté qu'il n'évoque nullement. Doit-on alors considérer sa réponse comme « réaliste » (puisqu'il procède à un arrondi) ou comme non réaliste (puisqu'il ne cerne pas le cœur de ce problème et « colle » – mathématiquement seulement – les restes de ses planches) ? La réponse « 4 planches » (3) traduit une « rupture » avec le « contrat didactique » classique (fournir comme réponse une donnée de l'énoncé étant rarement attendu en classe) et est possible en situation réelle (comme toute réponse proposant un nombre entier égal ou inférieur à 8). On pourrait alors penser qu'il est logique de la coder comme « RR ». Dans le dernier exemple (4), la démarche est peu compréhensible (l'élève retire-t-il 1 mètre parce qu'il voit qu'il y a des restes ?) et la réponse « 6 planches » n'est pas réaliste (même si elle est inférieure à 8) puisqu'elle découle de l'opération « 4x1,5=6 » qui correspond en réalité à la même logique que la « NR » conduisant à la réponse « 10 ». *A priori*, cette production est donc plutôt non réaliste. Un dernier exemple pour terminer et illustrer encore une autre difficulté rencontrée : comment considérer une réponse du type « *Pas clair ! On ne dit pas combien de mètre fait l'étagère* » ?

V. CONCLUSION PROVISOIRE ET DISCUSSION

Depuis de nombreuses années, des chercheurs s'intéressent aux réactions des élèves face à des problèmes qualifiés de *problématiques* dans la mesure où ils ne peuvent pas être résolus par la simple application d'une opération réalisée au départ des données de l'énoncé, du moins si l'on prend en considération des connaissances réalistes liées à la situation. Dans cet article, nous avons présenté les premières étapes d'une recherche visant à analyser l'impact d'*illustrations informationnelles* sur le caractère réaliste des réponses fournies par les élèves. Nous avons illustré les premiers résultats bruts au départ d'un problème classiquement utilisé dans ces études (les planches). Celui-ci présente la particularité de pouvoir être résolu par une réponse numérique précise, ce qui n'est pas le cas de la plupart des problèmes problématiques proposés dans ces études (voir exemples dans le tableau 1). On pourrait discuter longuement du choix des problèmes eux-mêmes ou des illustrations qui les accompagnent (pourquoi le menuisier commence-t-il par découper « un reste » plutôt qu'une planche de 1 m ?), mais ce n'est pas l'objet qui a retenu notre attention ici.

Cette analyse exploratoire ne nous permet évidemment pas encore d'apporter une conclusion quant à l'impact des *illustrations informationnelles*, mais les premiers résultats (non présentés ici), issus du codage brut, montrent une légère tendance en faveur de notre hypothèse. Ainsi, on note une faible augmentation des réponses réalistes attendues et une diminution des réponses non réalistes. Comme l'ont montré des études antérieures, quand elles sont utilisées face à des problèmes standards, les *illustrations informationnelles* complexifient la résolution du problème (Berends & Van Lieshout, 2009 ; Elia, 2009 ; Elia et al., 2007). Puisque notre objectif était d'empêcher les élèves de se précipiter dans la production d'une réponse non réaliste, nous pensions qu'elles pourraient s'avérer fonctionnelles face aux problèmes problématiques. Toutefois, la diminution des NR, qui entraîne une augmentation des réponses « autres » (AR) face à certains problèmes, pourrait aussi traduire un accroissement de la difficulté, engendré par la présence d'illustrations. Les élèves commettraient ainsi d'autres types d'erreurs que les NR attendues, mais sans pour autant fournir davantage de RR. Evidemment, il ne s'agit à ce stade que d'une hypothèse interprétative qui devra être vérifiée lors du recodage des réponses pour l'ensemble des problèmes.

Les éléments présentés ici sont centrés sur un problème spécifique, mais les difficultés de codage se sont révélées présentes pour tous les problèmes de l'étude. A ce stade, l'analyse exploratoire réalisée questionne la pertinence de vouloir classer toutes les réponses de façon dichotomique en « RR » *vs* « NR », comme c'est le cas dans les études antérieures. En nous centrant sur l'étape de codage, nous avons voulu mettre à jour la diversité des productions d'élèves et les difficultés d'interprétation de certaines réponses. Cette étape constitue un enjeu central des recherches visant à évaluer l'impact de certaines variables sur les réponses produites par les élèves dans un test, enjeu qui nous semble pourtant insuffisamment développé dans de nombreux travaux.

REFERENCES

Baruk, S. (1985). *L'âge du capitaine : de l'erreur en mathématiques*. Paris : Seuil.

Berends, I.E., & van Lieshout, E.C.D.M. (2009). The effect of illustrations in arithmetic problem-solving: Effects of increased cognitive load. *Learning and Instruction, 19,* 345-353.

Brousseau, G. (1990). Le contrat didactique : Le milieu. *Recherches en Didactique des mathématiques, 9,* 308-336.

Depaepe, F., De Corte, E., & Verschaffel, L. (2015). Students' non-realistic mathematical modeling as a drawback of teachers' beliefs about and approaches to word problem solving. In B. Pepin & B. Roesken-Winter (Eds.). *From beliefs to dynamic affect systems in mathematics education.* Springer: Advance in mathematics education.

Dewolf, T., Van Dooren, W., Ev Cimen, E., & Verschaffel, L. (2014). The impact of illustrations and warnings on solving mathematical word problems realistically. *The Journal of Experimental Education, 82*(1), 103-120.

Dewolf, T., Van Dooren, W., & Verschaffel, L. (2017). Can visual aids in representational illustrations help pupils to solve mathematical world problems more reastically? *European Journal of Psychology of Education, 32,* 335-351.

Elia, I. (2009) L'utilisation d'images dans la résolution de problèmes additifs: quel type d'images et quel rôle ? *Annales de Didactiques et des Sciences cognitives, 14*, 5-29.

Elia, I., Gagatsis, A., & Demetriou, A. (2007). The effects of different modes of representations on the solution of one-step additive problems. *Learning and Instruction, 17,* 658-672.

Elia, I., & Philippou, G. (2004). The functions of pictures in problem solving. In M. J. Hoines & A. B. Fuglestad (Eds.), *Proceedings of the 28th conference of the international group for the psychology of mathematics education* (Vol. 2, pp. 327–334). Bergen, Norway: University College.

Mellone, M., Verschaffel, L., & Van Dooren, W. (2017). The effect of rewording and dyadic interactions on realistic reasoning in solving word problems. *The Journal of Mathematical Behavior, 46*, 1-12.

Van Dooren, W., Verschaffel, L., Greer, B., De Bock, D., & Crahay, M. (2015). La modélisation et la résolution de problèmes d'application. In M. Crahay & M. Dutrevis (Eds), *Psychologie des apprentissages scolaires* (pp. 199-220). Bruxelles: De Boeck.

Verschaffel, L., & De Corte, E. (2008). La modélisation et la résolution des problèmes d'application : de l'analyse à l'utilisation efficace. In M. Crahay, L. Verschaffel, E. De Corte & J. Grégoire (Eds.). *Enseignement et apprentissage des mathématiques. Que disent les recherches psychopédagogiques ?* (153-176). Bruxelles : De Boeck.

Verschaffel, L., De Corte, E., & Lasure, S. (1994). Realistic considerations in mathematical modelling of school arithmetic word problems. *Learning and Instruction, 4,* 273-294.

Verschaffel, L., Greer, B., & De Corte, E. (2000). *Making sense of word problems*. Lisse, The Netherlands : Swets & Zeitlinger.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation and autonomy in mathematics. *Journal for Research in Mathematics Education, 27*(4), 458-477.