

# Validité d'épreuves d'évaluation mettant en jeu des illustrations au CP : mise en œuvre d'une méthodologie d'analyse de quelques épreuves

Christophe Blanc - [Christophe.Blanc@ac-paris.fr](mailto:Christophe.Blanc@ac-paris.fr)

Marc Vantourout - [marc.vantourout@parisdescartes.fr](mailto:marc.vantourout@parisdescartes.fr)

Sylvette Maury - [sylvette.maury@parisdescartes.fr](mailto:sylvette.maury@parisdescartes.fr)

Université Paris Descartes, Paris, France

**Pour citer cet article :** Blanc, C., Vantourout, M., & Maury, S. (2018). Validité d'épreuves d'évaluation mettant en jeu des illustrations au CP : mise en œuvre d'une méthodologie d'analyse de quelques épreuves. *Évaluer. Journal international de recherche en éducation et formation*, 4(1), 21-36.

## Résumé

Cet article vise à montrer en quoi les illustrations accompagnant des épreuves d'évaluation peuvent questionner leur validité. Il s'organise autour de l'analyse de quelques épreuves proposées par des enseignants de CP (Cours Préparatoire, 1st grade) à leurs élèves, en mathématiques et en français. Le cadre d'analyse utilisé est celui des approches psycho-didactiques des évaluations (APDE ; Vantourout et Goasdoué, 2014), dans lequel l'analyse conceptuelle de l'épreuve est complétée par la prise en compte de certains éléments du contexte de la tâche proposée, conduisant ainsi à une meilleure prédiction de l'activité des élèves. Il apparaît que les illustrations, même lorsqu'elles sont plutôt a priori dotées d'une simple fonction « motivationnelle », sont en fait susceptibles d'intervenir au niveau du processus de redéfinition de la tâche. Elles sont ainsi une source potentielle d'écart entre la tâche attendue par l'enseignant et celles effectivement réalisées par les élèves, ce qui conduit à remettre en cause la validité de ces épreuves, au sens des APDE.

## Mots-clés :

Évaluation, validité psycho-didactique (VPD), illustrations, cours préparatoire (CP).

## Abstract

This article is about the validity analysis of assessment tasks that include illustrations. Those tasks are used by 1st Grade teachers (for their pupils) in Mathematics and French. Our analysis is based on the assessment psycho-didactic approaches (APDA ; Vantourout & Goasdoué, 2014). Some context elements of the tasks are taken into account in order to reinforce the conceptual analysis of the task. This leads to a better prediction of the pupils activity. It seems that illustrations participate in the task redefinition process even though their function is rather only "motivation orientated". Those illustrations are a gap potential source between the task that the teacher expects and the tasks the pupils indeed realize. This leads to challenge the validity of those APDA tasks.

## Keywords

Assessment, psycho-didactic validity (PDV), illustrations, first grade.

## 1. Introduction

Dans cet article nous nous intéressons à la validité d'une épreuve<sup>1</sup> en tant qu'outil d'évaluation, c'est-à-dire à sa capacité à évaluer effectivement la compétence visée par celui qui prescrit l'épreuve : l'enseignant. Nous le faisons dans le cadre d'une approche psycho-didactique des évaluations (APDE) qui nous conduit à nous prononcer sur la validité psycho-didactique d'une épreuve (VPD) (Vantourout & Goasdoué, 2014). Avec ce type d'approche qui conjugue les apports des didactiques disciplinaires et de l'ergonomie cognitive, on complète l'analyse conceptuelle, propre aux didactiques, par une analyse de la tâche, en s'interrogeant sur ce que la tâche et ses spécificités sont susceptibles d'engendrer quant à l'activité de réponse de l'élève évalué. Des travaux (Grapin, 2015 ; Vantourout & Maury, 2017) attestent en effet que cette approche conduit à une meilleure appréciation de la qualité de l'évaluation que ne le fait la simple analyse conceptuelle, dans le sens où elle permet une meilleure prédiction de l'activité réelle de l'élève. Nous référant au modèle d'analyse psychologique de l'activité de Leplat (1997), revisité par Rogalski<sup>2</sup> (2003), nous analyserons ainsi dans l'article, sur quelques exemples, l'écart éventuel entre la tâche attendue par l'enseignant, la tâche prescrite et la tâche telle qu'elle pourrait être redéfinie par l'élève à des fins de réalisation.

Au niveau du CP, auquel nous nous intéressons ici, les élèves (6-7 ans) se caractérisent plus que tous les autres dans l'enseignement obligatoire par le fait qu'ils ne sont pas des lecteurs experts. Cela constitue une contrainte importante pour les enseignants lorsqu'ils définissent les épreuves d'évaluation. Une autre contrainte qui s'impose à eux consiste à évaluer tous les élèves d'une classe sur une période réduite afin de renseigner les livrets scolaires<sup>3</sup>, ce qui les conduit généralement à recourir massivement à des épreuves écrites, individuelles, passées simultanément par l'ensemble de la classe. C'est ainsi que les enseignants ont très largement recours aux illustrations<sup>4</sup> dont ils postulent sans doute qu'elles concourent à rendre intelligibles les tâches, mais aussi qu'elles sont de nature à soutenir l'engagement dans la tâche et la réalisation autonome de celle-ci.

Il existe, depuis longtemps, de nombreux travaux interrogeant les fonctions didactiques ou cognitives des illustrations dans l'enseignement ou l'apprentissage de la lecture, des mathématiques et des sciences (Baillé & Maury, 1993 ; Monnier, 2003 ; Peraya, 1995 ; Vézin, 1986). Le rôle bénéfique des illustrations dans la résolution de problèmes ainsi que dans la mémorisation ou la compréhension de textes y est régulièrement souligné (Gyselinck, 1996 ; Monnier, 2003). Cependant, si la fonction motivationnelle des illustrations est rarement mise en doute, en particulier à l'école élémentaire, leur prétendu rôle facilitateur est souvent discuté comme le mentionne Vézin (1986, p. 111) : « s'il est certain que l'illustration, de par son caractère concret motive l'élève, elle peut aussi détourner l'attention et entraver l'apprentissage ». Des travaux plus récents (Berends & van

---

<sup>1</sup> Le terme « épreuve » désigne dans cet article un « exercice » d'évaluation, comprenant une ou un nombre limité de questions. Il ne désigne pas un test, constitué d'un ensemble d'exercices.

<sup>2</sup> Leplat (1997) distingue, du côté du prescripteur, la « tâche à réaliser » et la « tâche prescrite », du côté du réalisateur, la « tâche redéfinie » et la « tâche effective ». Rogalski (2003), qui reprend cette typologie pour analyser l'activité des enseignants, préfère parler de « tâche attendue » plutôt que de « tâche à réaliser ». Nous la rejoignons, la tâche attendue étant « le contenu réel des attentes du prescripteur » (2003, p. 350).

<sup>3</sup> Le livret scolaire de l'école primaire, en vigueur lorsque nous avons collecté les épreuves analysées dans cet article, permet d'attester progressivement des compétences et des connaissances acquises par l'élève dans les divers domaines enseignés. A la rentrée scolaire 2016, il a été remplacé par « Le nouveau livret scolaire unique de l'école et du collège ».

<sup>4</sup> Comme le font d'autres auteurs (Gyselinck, 1996 ; Peraya, 1995) nous utiliserons le terme « illustration » comme un terme générique recouvrant des dessins, des graphiques, des images plus ou moins schématiques qui accompagnent un texte, une consigne, un énoncé...

Lieshout, 2007 ; Elia, 2009 ; Elia, Gagatsis, & Demetriou, 2007) étudiant l'effet de divers types d'illustrations sur la résolution de problèmes additifs, montrent d'ailleurs que certains types d'illustrations peuvent avoir un effet préjudiciable sur les performances des élèves et, s'agissant de problèmes arithmétiques « non standards »<sup>5</sup>, une série de recherches expérimentales (Dewolf, Van Dooren, Ev Cimen, & Verschaffel, 2014 ; Reuter, Schnotz, & Rasch, 2015) mettent en évidence que l'effet facilitateur (attendu) des illustrations n'est généralement pas avéré. Enfin, des travaux menés au sein de notre équipe attestent que le rôle des représentations externes n'est pas aussi transparent qu'on peut le supposer *a priori* (Maury, 2002 ; Rudat & Maury, 2009).

On peut rendre compte des difficultés rencontrées par les élèves, comme nous le ferons, en référence aux modèles classiques de la psychologie cognitive auxquels didacticiens et psychologues ont régulièrement recours pour décrire la compréhension en lecture (pour une synthèse, Blanc & Brouillet, 2003 ; Gyselinck, 1996) et en mathématiques, notamment lors de la résolution de problèmes (Nathan, Kintsch & Young, 1992 ; Thevenot, Coquin & Verschaffel, 2006). Ces modèles ont en commun la place accordée à la construction par le sujet d'une représentation, selon divers processus, notamment par la construction d'un modèle particularisé de situation ou par la particularisation d'un schéma (Richard, 2004). Ainsi, d'après Devidal, Fayol et Barrouillet (1997, p. 9), qu'il s'agisse d'un texte littéraire ou d'un énoncé de problème, « comprendre ce qu'on lit c'est construire un « modèle mental »<sup>6</sup> (Johnson-Laird, 1983, 1993) ou un « modèle de situation », (Kintsch, 1979) de ce qui est relaté dans le texte ». Un tel modèle consiste en l'intégration des informations fournies par le texte aux connaissances antérieures du lecteur, notamment aux connaissances sémantiques sur les situations évoquées, et/ou aux schémas dont celui-ci peut disposer (Julo, 1995 ; Richard, 2004 ; Thevenot *et al.*, 2006).

C'est à une analyse qualitative d'épreuves, mettant en jeu des illustrations<sup>7</sup> utilisées par des enseignants de CP dans le cadre de l'évaluation interne, que le présent article est consacré. Dans la perspective qui est la nôtre ici – celle de la validité dans le cadre des APDE – la question première n'est pas celle de savoir si les illustrations facilitent ou non la réussite. Il s'agit d'étudier, dans des cas précis, en quoi une illustration est susceptible (ou non) d'invalider une épreuve en tant qu'outil d'évaluation, c'est-à-dire dans sa capacité à évaluer effectivement la compétence ou la connaissance visée par l'enseignant. Les épreuves examinées dans l'article ont été soumises par des enseignants à leurs élèves à des fins d'évaluation et servent de support aux jugements évaluatifs portés dans les livrets scolaires. Elles sont extraites d'un travail plus ample (Blanc, 2017) dans lequel l'auteur s'intéresse aux divers types de validité d'épreuves ou de tests d'évaluation<sup>8</sup> utilisés par 16 enseignants de CP<sup>9</sup>, concernant les sous-domaines du français et des mathématiques, respectivement intitulés « lecture » et « nombre et calcul ». L'objectif ici n'est pas de rapporter l'ensemble des résultats mais de montrer concrètement, sur la base d'une méthode originale d'analyse

<sup>5</sup> Problèmes requérant un raisonnement « réaliste » (*realistic word problems*) chez Dewolf *et al.* (2014), problèmes « non-routiniers » (*non-routine word problems*) chez Reuter *et al.* (2015).

<sup>6</sup> Modèle mental et modèle de situation sont généralement utilisés comme des synonymes (Gyselinck, 1996) ; nous retiendrons dans la suite le terme de modèle de situation.

<sup>7</sup> Nous entendons par là une épreuve dont l'énoncé contient une illustration ou (ce sera le cas de l'épreuve 2 examinée dans le paragraphe 3) la demande explicite, adressée à l'élève, de produire un dessin.

<sup>8</sup> L'auteur s'intéresse à la validité dite « curriculaire » de tests d'évaluation, à leur validité didactique ainsi qu'à la validité psycho-didactique d'épreuves, à travers notamment l'examen de l'ensemble des épreuves d'évaluation utilisées par les enseignants concernés aux 1<sup>er</sup> et 3<sup>e</sup> trimestres (soit 754 épreuves). Il montre que la validité – quel que soit le type envisagé – n'est pas, dans bien des cas, assurée.

<sup>9</sup> Cet « échantillon » est composé d'enseignants exerçant dans 13 écoles différentes réparties sur 6 circonscriptions scolaires. Bien qu'assez diversifié, il n'a cependant pas la prétention d'être représentatif, de quelque manière que ce soit, de la population des enseignants de CP.

de la validité, et en nous appuyant sur quelques épreuves (quatre épreuves émanant de quatre enseignants différents), ce que peut être l'effet du recours aux illustrations sur cette validité. Chacune des épreuves a été affectée, par l'enseignant qui l'a proposée à ses élèves, à un item<sup>10</sup> du livret scolaire qu'il utilise, cet item constituant une manière de décliner la compétence ou la connaissance dont l'épreuve est censée contribuer à évaluer l'acquisition : ainsi, « résoudre un problème simple à une opération », « repérer graphies et syllabes dans un mot » sont deux exemples d'items tels qu'on peut les rencontrer dans les livrets scolaires. Pour chaque épreuve examinée, afin de contribuer à l'identification de la tâche attendue par le prescripteur, nous indiquerons l'item du livret auquel l'enseignant concerné l'a référée. Les enseignants nous ont transmis l'énoncé de chaque épreuve par le biais d'une copie d'élève que nous analyserons, afin de mieux éclairer notre propos.

En définitive, le questionnement qui est au centre de cet article peut se formuler de la manière suivante : la présence ou la demande de production d'illustrations (cf. note 7) dans une épreuve d'évaluation, en intervenant notamment dans le processus d'élaboration du modèle de situation, pourrait avoir pour conséquence de générer un écart entre la tâche attendue par le prescripteur et la tâche effectivement réalisée par l'élève, et donc remettre en cause la validité de l'épreuve.

Conformément à l'objectif de l'article – montrer les effets que peuvent avoir des illustrations sur la validité d'épreuves d'évaluation – nous avons retenu, à dessein, quatre épreuves problématiques. Celles-ci vont nous permettre de présenter, dans le cadre que nous nous sommes fixé (celui de la VPD), les diverses facettes de notre questionnement et la méthode d'analyse. Pour témoigner du caractère généralisable de cette méthode, les quatre épreuves sont « diversifiées » dans le sens où elles portent sur divers points des sous-domaines concernés et correspondent à des niveaux de compétence a priori différents<sup>11</sup> : l'épreuve 1 se réfère aux décompositions additives ; l'épreuve 2 à la résolution de problèmes ; l'épreuve 3 à la compréhension de textes et l'épreuve 4 aux habiletés phonologiques. Les épreuves 1 et 4 mettent en jeu des compétences de base alors que les épreuves 2 et 3 renvoient à des compétences complexes (cf. note 11).

---

<sup>10</sup> Un item du livret scolaire correspond à une ligne où est désignée une « compétence » ou une « connaissance ». En général, dans le champ de l'évaluation, le terme « item » désigne un exercice ou une question. Ce ne sera jamais le cas dans cet article où nous réservons ce terme à l'usage qui en est fait dans le livret.

<sup>11</sup> On distingue dans la littérature, depuis longtemps (Tourneur, 1979), deux principaux niveaux de compétences que Perrenoud (2011) clarifie ainsi : les habiletés (premier niveau) sont des ressources au service de compétences plus globales (deuxième niveau) renvoyant à des situations complexes (p. 66). Aux habiletés (ou encore compétences « de base » ou « élémentaires ») et aux compétences (dites parfois « complexes ») sont respectivement associées les tâches « simples » et « complexes ». Ces distinctions sont reprises dans des documents institutionnels, notamment dans l'un deux, en 2011, où il est indiqué que « les tâches simples incitent davantage à des reproductions de procédures laissant peu d'initiative à l'élève alors que les tâches complexes permettent une stratégie de résolution propre à chaque élève. » ([http://media.eduscol.education.fr/file/DNB/89/2/socle-C3-vade\\_mecum\\_166892.pdf](http://media.eduscol.education.fr/file/DNB/89/2/socle-C3-vade_mecum_166892.pdf)).

## 2. Analyse de la validité de l'épreuve 1

Un enseignant utilise l'épreuve 1 (fig. 1) pour évaluer conjointement deux items du livret scolaire : « savoir écrire et nommer les nombres » et « produire les décompositions additives des nombres inférieurs à 20 ».



**Figure 1.** Épreuve 1 avec la réponse d'un élève (la correction, écriture correcte de « 3 », a été effectuée par l'enseignant)

Au vu des items du livret auxquels l'enseignant réfère l'épreuve, la tâche attendue est une tâche de décomposition additive, mobilisant des connaissances relatives à la désignation du nombre (lecture de « 5 », écriture de « 3 ») : l'élève devrait ainsi déterminer le complément de 5 à 8 (par calcul, ou par sur-comptage<sup>12</sup>), tâche qui peut être réalisée sans recourir aux informations apportées par l'illustration, tout se passant alors comme si celle-ci occupait une simple fonction motivationnelle<sup>13</sup>.

Intéressons-nous au rôle potentiel joué par l'illustration pour un élève qui réaliserait la double consigne (Observe le dessin. Complète l'addition) « à la lettre », en commençant par observer l'illustration. L'élève pourrait dénombrer les collections d'enfants représentés, « les cinq personnages assis » et « les trois personnages debout », puis traduire le résultat de ces dénombrements dans la numération chiffrée. Mais il pourrait aussi bien dénombrer les sièges, comme la remarque inscrite dans la bulle, « Il y a encore 3 places ici », pourrait l'inciter à le faire (les sous-catégories étant alors libres *vs.* occupées). Une difficulté pour un élève qui procéderait ainsi résiderait dans l'élaboration du modèle de situation : il peut en particulier y avoir un doute sur la nature des informations pertinentes à retenir (les enfants ? les sièges ?), même si dans les deux cas la compétence à l'œuvre – celle du dénombrement des éléments d'une collection – conduit au même résultat. La double consigne suggère ensuite l'engagement dans une activité de modélisation qui s'exprime mathématiquement par «  $5+3=8$  ». En fait, la compréhension (par un élève) de la situation évoquée par l'illustration pourrait s'apparenter à un récit du type : « il y a des enfants (ou des chaises) dans une salle, cinq sont assis (ou occupées) et trois sont debout (ou libres), ils (elles) sont huit en tout », et de là, on peut écrire  $5+3=8$ . En réalité, si l'on considère la

<sup>12</sup> L'élève peut « sur-compter » à partir de 5 en énonçant la suite des mots-nombres jusqu'à obtenir 8 et en dénombrant parallèlement le nombre de mots-nombres prononcés au moyen d'un autre compteur (les doigts par exemple).

<sup>13</sup> Parmi les auteurs qui s'intéressent aux questions cognitives ou didactiques, il existe diverses propositions de classification des fonctions des illustrations (Carney & Levin, 2002 ; Denis, 1989 ; Elia & Philippou, 2004 ; Gyselinck, 1996 ; Monnier, 2003 ; Peraya, 1995 ; Vezin, 1986). Celles-ci sont souvent liées à la discipline et aux intérêts particuliers de l'auteur et aucune ne convient parfaitement pour traiter les situations auxquelles nous nous intéressons. Dans le cas précis qui nous occupe, certains auteurs parleront de fonction « illustrative » ou « esthétique » ou « évocatrice » ou encore « représentative », voire « compensatrice » ou bien, comme le fait Vézin (1986), de fonction motivationnelle ; c'est en effet ainsi qu'elle qualifie ce type d'illustrations qui « attire l'attention du lecteur et rend le texte plus intéressant à lire ». C'est ce dernier terme que nous retiendrons ici.

redondance des informations au sein de l'illustration (données « enfants » et « sièges »), la tâche prescrite s'apparente à une forme particulière de tâche complexe.

En définitive, si nous considérons les différents éléments constituant cette épreuve et leur complémentarité, l'illustration pourrait remplir, pour certains élèves, une fonction de « complexification » de la tâche attendue. En effet, savoir écrire le nombre 3 en écriture chiffrée et connaître une décomposition additive de 8 renvoient en fin de CP à ce que l'on appelle des « habiletés » ou des « compétences de base », celles-ci nécessitant pour leur évaluation des tâches dites « simples ». Or, comme nous l'avons montré, l'illustration et la double consigne conduisent potentiellement à prescrire, au lieu de ce qui pourrait être une tâche simple<sup>14</sup>, une tâche bien plus « complexe ».

En admettant que ces difficultés soient surmontées, la réalisation de cette tâche ne nécessiterait pas la production d'une décomposition additive : reporter « 3 » (3 élèves debout ou 3 chaises vides) dans l'addition « à trou » suffit en effet pour réussir, sans avoir eu à construire que l'addition de 5 et de 3 produisait 8 ou que 8 se décomposait en additionnant 5 et 3. Autrement dit, une fonction de l'illustration résiderait ici dans l'apport d'information (3 élèves ou 3 chaises) et les tâches redéfinies et effectives renverraient alors au dénombrement. En définitive, la VPD de l'épreuve n'est pas avérée, la présence de l'illustration pouvant alors favoriser l'engagement de l'élève dans un dénombrement plutôt que dans une décomposition additive, sans que la qualité de la réponse produite (le report de « 3 ») permette de savoir si l'élève sait ou a su décomposer 8 en 5+3.

### 3. Analyse de la validité de l'épreuve 2

Un enseignant, en fin de CP, utilise l'épreuve 2 (fig. 2) pour évaluer l'item de son livret, correspondant à une compétence complexe, « Résoudre un problème simple à une opération ». Cette épreuve diffère de la précédente dans le sens où aucune illustration ne figure dans son énoncé ; en revanche, la production d'un dessin par l'élève est exigée, comme l'indique la consigne : « dessine les fruits puis écris une opération ».

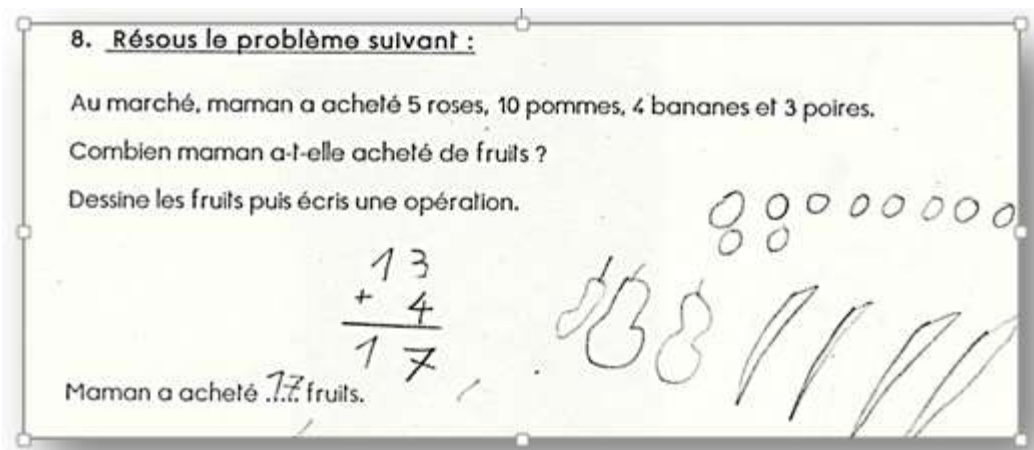


Figure 2. Épreuve 2 avec réponse d'un élève

<sup>14</sup> Par exemple : « complète l'opération à trou  $5 + \dots = 8$  ».

Le problème présente une donnée « surabondante » (5 roses) ; l'élève doit ainsi effectuer, lors de l'extraction des données pertinentes, une tâche de catégorisation : il faut qu'il reconnaisse les différents éléments de l'énoncé comme appartenant, ou n'appartenant pas, à la catégorie des fruits. Il est cependant peu probable que les élèves, en fin de CP, échouent à cette tâche de catégorisation, tous les éléments pouvant être considérés comme familiers. La donnée surabondante peut également intervenir lors de l'élaboration du modèle de situation et certains élèves peuvent être tentés, par un effet du contrat didactique (Brousseau, 1988), d'utiliser tous les nombres figurant dans l'énoncé. Cela constitue une source potentielle de difficulté commune à tous les problèmes à donnée(s) surabondante(s), problèmes que l'on hésite d'ailleurs généralement à qualifier de « simples »<sup>15</sup>.

Si l'on fait abstraction de la donnée surabondante ainsi que de la prescription de dessiner les fruits, le problème se présente comme un problème classique de composition de mesures (Vergnaud, 1981)<sup>16</sup> et s'inscrit en cohérence avec l'item auquel l'épreuve est référée ; la tâche attendue par l'enseignant, en fin de CP, pourrait être alors une modélisation mathématique de la situation par une opération additive. Tout élève qui traiterait directement le problème comme tel, serait ainsi amené à écrire l'opération  $10+4+3=17$  (en ligne, voire même en colonne en fin de CP), en calculant  $10+4$  : « 14 »,  $14+3$  : « 17 », et à conclure « Maman a acheté 17 fruits ».

Quel effet pourrait avoir la nécessité de produire en premier lieu un dessin comme le prescrit la consigne ? L'élève qui procéderait ainsi « transformerait » le texte, dans ce premier temps, sous une forme (le dessin produit) qui lui permettrait de se focaliser sur l'information critique et « d'oublier » l'information surabondante, ce qui pourrait agir directement sur l'élaboration du modèle de situation, l'illustration remplissant alors une fonction « modélisante » : le problème à donnée surabondante se verrait ainsi transformé en un problème « simple » de composition de mesures. Mais en revanche, le risque serait grand que, ce faisant, pour répondre à la question posée « combien maman a-t-elle acheté de fruits ? », l'élève réalise un simple dénombrement de la collection représentée dans son dessin, manifestant ainsi une redéfinition de la tâche attendue en une tâche de dénombrement.

Examinons la réponse que nous avons recueillie (fig. 3) : l'élève a dessiné tous les fruits conformément à la prescription de l'énoncé. Mais il est intéressant de remarquer qu'il n'y a pas congruence entre l'ordre des données de l'énoncé (les pommes, les bananes et enfin les poires) et celui selon lequel elles apparaissent dans son dessin : une lecture de celui-ci de gauche à droite fait en effet apparaître, dans l'ordre, les pommes puis les poires, enfin les bananes<sup>17</sup>. L'apparition du nombre « 13 » dans l'opération posée par l'élève («  $13+4=17$  ») indiquerait alors que, selon toute vraisemblance et comme nous le mentionnions plus haut, il s'est appuyé, pour répondre, sur son dessin (en déterminant en premier le nombre d'éléments « pommes, poires ») et non directement sur l'énoncé. Ce faisant, il a pu obtenir « 13 » par le calcul («  $10+3=13$  ») ou, le dessin l'y engageant, par un comptage ou un sur-comptage (cf. note 9) en énumérant « 11, 12, 13 ». Il a pu réitérer la procédure avec les 4 bananes à partir de 13 : « 14, 15, 16, 17 » et enfin reporter le dernier mot-nombre prononcé

<sup>15</sup> Les difficultés liées à la résolution des problèmes dits « non-standard » (et en particulier aux problèmes à données surabondantes) ont fait l'objet de nombreux travaux (pour une revue voir Jimenez & Verschaffel, 2014). Les interprétations de ces difficultés se font classiquement en invoquant des effets du contrat didactique (Brousseau, 1988) ou des normes socio-mathématiques (Yackel & Cobb, 1996).

<sup>16</sup> On trouve parfois ce type de problèmes sous le nom de « problème de combinaison » utilisé par Riley, Greeno & Heller (1983) dans la classification des problèmes additifs qu'ils proposent (voisine de celle de Vergnaud, 1981).

<sup>17</sup> Il est vraisemblable que l'élève, après avoir dessiné les pommes puis les bananes, ait dessiné les poires dans le seul espace disponible sur la feuille, c'est à dire à gauche des bananes.



dans la phrase « à trou ». L'opération pourrait alors avoir été posée par cet élève *a posteriori* du processus de résolution du problème afin de respecter toutes les prescriptions de l'énoncé, parmi lesquelles figure « écris une opération ». Il est ainsi possible que l'élève ait utilisé une procédure de dénombrement pour résoudre le problème et reconstruit *a posteriori* une réponse qui convienne à la demande.

En définitive on peut contester la validité de l'épreuve, l'exigence d'un dessin dans la prescription de l'énoncé étant susceptible de conduire l'élève à la redéfinition de la tâche en une tâche de dénombrement d'une collection, alors que l'attente de l'enseignant, telle que formulée dans la consigne, nécessite la réalisation d'un calcul via une opération.


#### 4. Analyse de la validité de l'épreuve 3

Un enseignant utilise l'épreuve 3 (fig. 3) pour évaluer l'item : « Je comprends un texte ».

**6) Donner, après lecture, des renseignements sur un texte** 10 / 10

Consigne : Observe la couverture du livre. Lis le texte. Coche les bonnes réponses.

**Le petit Chaperon rouge**  
D'après Jacob Grimm



Il était une fois une petite fille qu'on appelait le petit Chaperon rouge.

Un jour, en allant chez sa grand-mère, elle rencontra le loup.

Vif et rusé, l'animal réussit à dévorer la vieille femme, puis la petite fille.

Heureusement, un bûcheron arriva et tua le monstre. Il délivra les prisonnières.

Elles vécurent alors très heureuses.

**A – Ce texte est :**

une recette

un conte

une lettre

**B – Son titre est :**

La petite fille aux allumettes

Le petit Chaperon rouge

Le vilain petit canard

**C – L'auteur de l'histoire s'appelle :**

Jacob Grimm

Hans Andersen

Charles Perrault

**D – C'est l'histoire :**

de Cendrillon qui épouse le prince.

du Chaperon rouge qui voit le loup

de Blanche Neige et 7 nains.

**E – Comment est le loup ?**

maladroit

rusé

peureux

**F – La petite fille rencontra le loup :**

en allant chez sa tante.

en allant chez le bûcheron.

en allant chez sa grand-mère.

**G – Qui tue l'animal ?** *Fais une phrase pour répondre*

le loup

Figure 3. Épreuve 3 avec les réponses d'un élève

Si l'on se réfère à la consigne de l'épreuve et à son item de référence, la tâche attendue vise l'évaluation d'une compétence complexe, la compréhension d'un texte, à partir d'un questionnaire portant sur le texte. Bien que l'utilisation de ce type d'épreuves ait été souvent critiquée (Goigoux & Cèbe, 2013 ; Vantourout & Maury, 2017), le questionnaire



reste l'outil que privilégient les enseignants pour évaluer la compréhension de textes (Fijalkov, 2003). Dans l'épreuve 3, le texte lui-même est accompagné d'un paratexte correspondant à la page de couverture, celle-ci comportant un titre, le nom de l'auteur et une illustration.

Outre l'observation de la couverture et la lecture du texte, la tâche prescrite comporte la lecture et le remplissage du questionnaire dont les six premières questions se présentent sous la forme de QCM (questions à choix multiples), alors que la septième est une question ouverte. L'analyse de l'illustration à elle seule permet à l'élève, qui connaîtrait le conte « Le Petit Chaperon rouge »<sup>18</sup>, de mobiliser des connaissances littéraires pour établir qu'il s'agit du personnage de ce conte et le reconnaître. Les réponses aux questions A, B et D sont alors immédiatement accessibles sans qu'il soit besoin de lire le texte : il suffit d'avoir reconnu le Petit Chaperon rouge sur le dessin (c'est le conte / intitulé le Petit Chaperon rouge / il raconte l'histoire du Petit Chaperon rouge). Les réponses aux questions E, F et G peuvent alors en découler : elles sont accessibles sur la base de la connaissance de ce conte (le loup est rusé / la petite fille rencontre le loup en allant chez sa grand-mère / le bucheron tue le loup). Ce type de questions – auxquelles on peut répondre correctement indépendamment de la lecture du texte, sur la base de connaissances déjà disponibles – est connu dans la littérature et décrit sous le nom de « *passage-independent questions* » (Keenan & Betjemann, 2006). Nous notons, par ailleurs, que la lecture du texte n'est pas nécessaire pour répondre à la question C, l'information requise figurant uniquement sur la page de couverture. Ainsi, dans le cas présent, pour un élève qui aurait une connaissance préalable du conte, les tâches redéfinies et effectives pourraient rendre compte - plutôt que de la compréhension du texte (dont la lecture peut être omise) - de sa capacité à gérer un questionnaire, notamment un QCM (lire les questions et les réponses proposées, identifier la bonne réponse, rejeter les distracteurs).

A l'inverse, un élève qui ne disposerait pas de connaissances sur ce conte serait conduit, pour répondre à l'ensemble des questions, à élaborer un modèle de la situation à partir de la lecture du texte, laquelle est difficile pour un élève de CP. L'élève devrait en outre être capable de parfaitement associer « le monstre » au « loup » et à « l'animal »<sup>19</sup>, « les prisonnières » à « la vieille femme et la petite fille », compétences qui ne sont pas nécessaires à celui qui répondrait aux questions sur la base de la seule considération de la couverture et de la connaissance du conte.

En définitive, la validité de l'épreuve peut être fortement contestée : nous avons mis en évidence que certains élèves pourraient répondre correctement à la grande majorité des questions du questionnaire sans avoir lu le texte de l'épreuve, alors que celle-ci est censée, si l'on se réfère à l'item retenu par l'enseignant, évaluer la compréhension de ce texte. Les connaissances intertextuelles des élèves peuvent totalement suffire et se substituer à la lecture. L'illustration, en permettant d'identifier dans sa globalité l'histoire à laquelle le texte se réfère, assurerait ici, non pas une simple fonction motivationnelle, mais une fonction de « représentation »<sup>20</sup>.

<sup>18</sup> Le conte du Petit Chaperon rouge fait partie de la liste de référence « La Littérature à l'école » pour l'école maternelle. Il est donc susceptible de faire partie du bagage de connaissances littéraires des élèves de CP.

<sup>19</sup> Dans la réponse recueillie (fig. 3), l'analyse des traces de l'activité de l'élève permet de considérer qu'il produit une réponse erronée à la seule question ouverte. Cette réponse, qui interroge quant au modèle de situation élaboré par l'élève, est difficile à interpréter. Il est cependant possible qu'il ne comprenne pas le lien entre « le loup » et ses substituts lexicaux « l'animal » et « le monstre » présents dans le texte, ou même qu'il comprenne mal la question en l'interprétant comme « qui est l'animal tué ».

<sup>20</sup> L'acception retenue ici est celle citée par Gyselinck (1996), empruntée par cet auteur à Levin, Anglin et Carney (1987) : l'illustration répète les contenus du texte.

## 5. Analyse de la validité de l'épreuve 4

Un enseignant utilise l'épreuve 4 (fig. 4) pour évaluer l'item « Connaître les correspondances entre les lettres et les sons dans les graphies simples et complexes ; reconnaître les syllabes et les mots à l'écrit et à l'oral ».

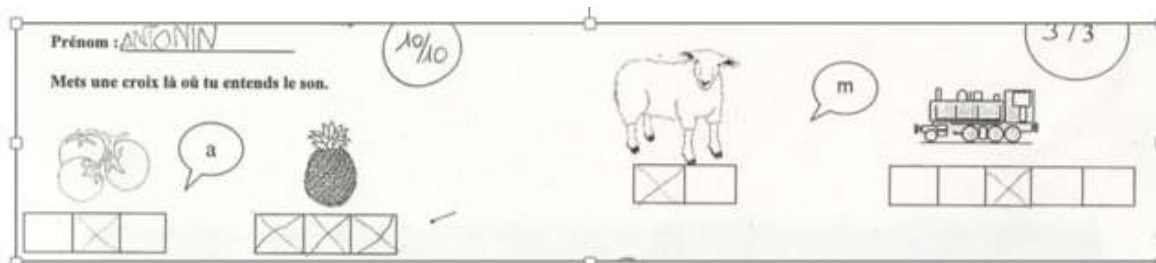


Figure 4. Épreuve 4 avec réponses d'un élève

L'épreuve 4 vise l'évaluation d'une compétence de base, plus précisément une habileté phonologique dans le domaine de la lecture<sup>21</sup>. Le recours aux illustrations est fréquent pour évaluer ce type d'habiletés (Blanc, 2017). Confronté à cette épreuve, qui se déroule entièrement à l'écrit, l'élève doit comprendre ce que l'on attend de lui. Pour cela, il doit construire un modèle de situation qui implique notamment l'identification des illustrations, laquelle repose sur la disponibilité et l'actualisation des schémas prototypiques correspondant à celles-ci chez les élèves évalués. Mais, les tâches phonologiques, contrairement aux précédentes, ne portent pas sur le sens véhiculé par ces illustrations une fois qu'elles ont été reconnues : l'une des difficultés de ces tâches, outre les habiletés en jeu, repose sur la capacité « à comparer ou à manipuler des sons indépendamment du sens des mots » (Ecalte & Magnan, 2002, p. 90), c'est-à-dire à « mettre entre parenthèses le sens ».

Intéressons-nous pour commencer à ce qui constitue, sur un plan phonologique, le « noyau » de cette épreuve, c'est-à-dire à la tâche attendue qui renvoie aux « savoir-faire », connaissances ou habiletés en jeu : localiser dans la chaîne sonore la syllabe dans laquelle se trouve le phonème cible<sup>22</sup>, ce qui sous-entend d'avoir préalablement effectué la segmentation syllabique du mot. Par exemple, si l'on veut repérer le phonème [m] dans « mouton » : [mutõ] se décompose en [mu-tõ] et [m] s'entend dans [mu].

Pour prescrire ce type de tâche, il n'est absolument pas nécessaire de recourir à des illustrations. Celles-ci ne servent qu'à présenter le matériel à partir duquel l'élève va exercer ses habiletés phonologiques. Le matériel (ou objet sur lequel va porter la tâche) est présenté sous une forme imagée, alors que la tâche attendue et sa réalisation requièrent une représentation sonore du mot. L'ensemble des options retenues pour cette épreuve (illustrations, archigraphème dans la bulle, cases à cocher) donne la possibilité d'évaluer des habiletés phonologiques en organisant une passation collective, autonome (chaque élève travaille à son rythme) et écrite. Si cette forme de passation présente effectivement un gain de temps, elle entraîne toutefois un coût supplémentaire à la charge de l'élève qui doit réaliser plusieurs tâches « ajoutées ». Celles-ci sont « implicites » sur le plan de la

<sup>21</sup> Depuis plusieurs années, des travaux soulignant la forte corrélation entre le niveau de conscience phonologique et l'acquisition de la lecture, ainsi que des propositions pédagogiques consacrées à la phonologie, sont largement diffusées en formation des enseignants (Goigoux, Cèbe & Paour, 2004). C'est vraisemblablement pour ces raisons qu'un bon nombre d'entre eux évalue des habiletés phonologiques, bien qu'elles ne figurent pas aux programmes de la classe de CP (en tout cas ceux de 2008, en vigueur lors du recueil de ces épreuves).

<sup>22</sup> Le son est représenté par un archigraphème placé dans une bulle. Il est implicite et laissé à la charge de l'élève qui doit comprendre que la bulle contenant la lettre « m » concerne les mots « mouton » et « locomotive ».

prescription et il est possible que certaines le restent lors de leur réalisation. Ces tâches ajoutées<sup>23</sup> sont notamment :

- reconnaître le mot représenté par l'illustration, c'est-à-dire reconnaître que l'un des dessins représente un mouton<sup>24</sup> ;
- puis, oraliser ce mot, être capable de travailler sur le signifiant sonore [mutõ] ;
- enfin, faire état de sa réponse par écrit en utilisant un codage.

Les modalités de réponses retenues, et les tâches « ajoutées » qu'elles impliquent, peuvent être, chez certains élèves, à l'origine d'un ensemble de difficultés. Il est possible que des enseignants recourant à ce type d'épreuves minimisent, voire ignorent, la charge de travail supplémentaire qui découle des choix opérés. Nous retrouvons l'un des axes d'analyse de la VPD, celui que nous qualifions d'« ergonomique » (Vantourout & Goasdoué, 2014).

Le dessin des carrés indique le nombre de syllabes que le mot contient et l'élève doit adapter sa réponse au cadre ainsi fourni. Le recours à cette représentation symbolique des syllabes est une source potentielle de difficultés, notamment celles liées à la distinction entre syllabes orales et syllabes écrites<sup>25</sup> : ici, par exemple, la réussite à la tâche concernant le mot « tomate » revient à cocher la case centrale ; or, l'élève peut adopter l'une ou l'autre des deux stratégies conduisant l'une à la bonne réponse, l'autre à une réponse erronée :

- soit il repère que le phonème [a] se situe dans la deuxième syllabe orale du mot et coche donc la deuxième case ;
- soit il repère que le phonème [a] se situe dans la dernière syllabe orale et coche alors la troisième case.

Le même type de remarque s'applique au mot « locomotive » qui contient quatre syllabes orales et cinq syllabes écrites. Avec une difficulté supplémentaire dans ce dernier cas : l'élève doit se donner oralement le mot « locomotive » mais il peut tout aussi bien attribuer la valeur « train » à l'illustration figurative et ne cocher aucune case.

La réalisation de la tâche mobilise a priori une compétence de segmentation syllabique qu'elle est censée évaluer (si l'on se fie à l'item de référence). Il est cependant possible de réussir, dans certains cas, sans mettre en œuvre cette compétence. C'est ce qui ressort de l'observation réalisée auprès d'élèves auxquels nous avons soumis la tâche (localiser le [m] de « mouton ») : l'un d'eux ne se souciait pas du nombre de syllabes et ânonnait plusieurs fois le phonème [m] pour former la suite [mmm..]. Ce faisant il cochait alors la première case, fournissant une réponse correcte mais qui ne traduisait pas la réalisation de la tâche attendue ; il n'avait pas segmenté le mot en deux syllabes et s'était appuyé, pour répondre, sur la seule attaque du mot.

En définitive, plusieurs facteurs peuvent mettre en question la VPD de cette épreuve. Certains ont directement trait à la segmentation syllabique et au schéma retenu pour la

---

<sup>23</sup> Bien sûr, il est tout à fait possible que les élèves d'une classe aient déjà rencontré auparavant des tâches analogues à celles qui servent à les évaluer. Les tâches ajoutées relèvent alors d'habitudes de classe pour la plupart des élèves, voire pour tous. Toutefois, il se peut également que, pour être évalués, des élèves soient confrontés à des tâches qu'ils n'ont jamais rencontrées auparavant, notamment en raison de l'utilisation d'évaluations « prêtes à l'emploi » que les enseignants trouvent sur des sites pédagogiques.

<sup>24</sup> La réussite à cette première tâche dépend, entre autres, des connaissances que l'élève a du monde qui l'entoure (a-t-il déjà rencontré de telles locomotives ?) et de la qualité des illustrations (les tomates sont-elles facilement identifiables en tant que tomates ? ne peut-on pas les confondre avec des pommes ?).

<sup>25</sup> Le mot « tomate » comprend deux syllabes orales [to-mat] et trois syllabes écrites TO/MA/TE. La segmentation en syllabes écrites, découpage que l'on fait à l'intérieur d'un mot, comme si tout se prononçait, est « indispensable aussi plus tard dans l'apprentissage de la lecture et de l'écriture » selon TFL (consulté à l'adresse <http://www.uvp5.univ-paris5.fr/TFL/ac/AffQpeRep.asp?CleFiche=P8-2>).

représenter : d'une part, alors que l'élève doit oraliser les mots, la schématisation de la segmentation syllabique imposée par l'énoncé correspond aux mots écrits ; d'autre part, comme nous l'illustrions dans le paragraphe précédent, l'activité de segmentation peut être, dans certains cas, contournée. Un autre facteur est directement lié à la fonction informative que remplissent les illustrations figuratives dans la communication des mots à traiter aux élèves. Le traitement sémantique et la traduction dans la langue naturelle qu'ils doivent opérer à partir du dessin figuratif ne sont pas forcément conformes à l'attente, ce qui risque fort d'échapper à l'enseignant puisque l'oralisation ne s'effectue pas en sa présence (l'élève oralise *a priori* « pour lui seul »).

## 6. Conclusion et discussion

Les épreuves que nous avons examinées ont été soumises par des enseignants à leurs élèves pour servir de support aux jugements évaluatifs portés dans les livrets scolaires. Ces livrets ont pour vocation d'être transmis aux familles et de matérialiser le suivi institutionnel de l'élève durant le cycle : l'enjeu de ces évaluations et des épreuves sur lesquelles elles sont fondées est donc, en terme de validité, particulièrement important.

Au CP, il est fait un grand usage d'illustrations dans les leçons, textes et divers énoncés destinés aux élèves, l'idée commune étant que la présence d'une illustration aurait une fonction facilitatrice. L'hypothèse que nous formulions était que, s'agissant d'épreuves à visée évaluative, l'intervention d'illustrations (figurant dans l'énoncé ou dont la production est sollicitée comme dans l'épreuve 3), au cours des processus d'élaboration du modèle de situation et de redéfinition de la tâche, pourrait au contraire, dans certains cas, engendrer un écart entre la tâche attendue et les tâches redéfinies et réalisées, et donc remettre en cause la validité de l'épreuve. C'est effectivement ce que nous mettons en évidence à travers l'analyse des quatre épreuves retenues, choisies à dessein problématiques. Il est clair cependant que, même si d'autres épreuves analysées dans Blanc (2017) conduisent aux mêmes conclusions, les données que nous avons examinées ne permettent pas vraiment de saisir l'ampleur du phénomène. Mais là n'était pas le but de l'article : en effet, au-delà de la mise en évidence du phénomène, l'un des objectifs de l'article était d'explicitier une méthode d'analyse qualitative originale de la validité, s'inscrivant dans une approche psycho-didactique des évaluations (APDE : Vantourout & Goasdoué, 2014), qui prenne en compte à la fois les aspects conceptuels de l'épreuve (comme on le fait classiquement en didactique) et les aspects ergonomiques de la tâche, en particulier ceux liés au recours à une illustration.

Ainsi, dans les épreuves de mathématiques examinées – qu'il s'agisse de décomposition additive (épreuve 1) ou de résolution de problème (épreuve 2) – nous mettons en évidence que la présence d'une illustration (ou l'exigence de sa production pour l'épreuve 2) est susceptible de conduire les élèves à redéfinir la tâche attendue en une tâche de dénombrement. Les fonctions effectivement assurées par les illustrations se situent alors au niveau de la modélisation de la situation (épreuve 2) ou de l'apport d'informations (épreuve 1 et 4), en lieu et place d'une simple fonction de motivation qui aurait pu leur être *a priori* dévolue par l'évaluateur. Dans l'épreuve de compréhension de texte (épreuve 3), l'illustration est susceptible d'assurer une fonction de représentation du texte concerné. Elle se substituerait au texte et détournerait ainsi de la tâche attendue par l'enseignant<sup>26</sup> (qui est une tâche de compréhension de texte). Enfin, dans l'épreuve 4 relative à la conscience

---

<sup>26</sup> Dans d'autres épreuves (non présentées ici), l'illustration est une représentation partielle du texte et ne se substitue qu'à une partie de celui-ci (Blanc, 2017).

phonologique, outre la présence d'un schéma modélisant le découpage en syllabes<sup>27</sup>, l'illustration figurative est porteuse d'informations indispensables à la résolution de la tâche, qui sont à traduire dans le registre de la langue naturelle et dont le traitement sémantique – même lorsqu'il est pertinent – n'assure pas une traduction conforme à l'attente et nécessaire à la réussite.

Lorsque l'on interroge des enseignants de CP, ils disent en général que les illustrations sont là pour motiver, agrémente, décorer, etc. Et il est vrai que, dans la majorité des épreuves analysées, l'illustration n'est pas indispensable à la résolution de la tâche (sauf pour l'épreuve 4, parce que l'on vise une passation autonome), les informations étant toutes présentes dans le texte ou l'énoncé verbal. Mais nous avons vu que la fonction effective qu'elles remplissent risque d'être toute autre que la fonction de motivation qui peut leur être prêtée *a priori*. La conclusion la plus saillante est que la validité des épreuves, dans les cas examinés, n'est pas assurée, ce qui conduit à s'interroger sur la validité des jugements évaluatifs prononcés, vis-à-vis de certains élèves, suite aux passations de ces épreuves.

Il est à noter que, pour conduire notre analyse, aucune des typologies d'illustrations que nous avons rencontrées dans la littérature ne s'est trouvée vraiment adaptée, comme nous le soulignons déjà dans la note 13. Il est vrai que notre travail n'avait pas pour but de tester le domaine de validité d'une typologie existante, les typologies proposées étant souvent liées à la discipline scolaire à laquelle l'auteur s'intéresse, alors que notre travail concerne à la fois les mathématiques et le français. Il est vrai également que notre travail n'avait pas pour objectif de produire une nouvelle typologie : ainsi, même si nous avons été amenés à distinguer certaines « fonctions » des illustrations, nous n'avons pas cherché à les traduire au sein d'une nouvelle proposition de typologie, le nombre d'épreuves examinées dans l'article étant de toute façon trop restreint pour avoir une telle ambition. Mais il est clair que si l'on se place dans la perspective d'aider les enseignants à choisir les illustrations qui accompagnent les épreuves d'évaluation qu'ils donnent à leurs élèves, la catégorisation des illustrations au sein d'une typologie constitue une piste de travail intéressante que nous nous proposons d'exploiter dans un travail expérimental.

Plus généralement, de nombreux enseignants sont démunis face aux prescriptions institutionnelles lorsqu'ils doivent évaluer leurs élèves, notamment en France (MENJVA-MESR, 2011 ; Gonthier-Maurin, 2012). Une entrée par l'analyse des tâches d'évaluation, en lien avec les didactiques disciplinaires (Bain, 2002 ; Roditi & Salle, 2015 ; Vantourout & Goasdoué, 2014), ainsi que des propositions d'ingénieries évaluatives courtes (Vantourout & Maury, 2017) pourraient constituer des pistes pour les former et les aider dans un domaine où les carences sont importantes. D'autant plus que ces carences ne concernent pas seulement les enseignants et les évaluations internes, mais aussi des évaluations nationales externes comme il a encore été possible de s'en rendre compte en septembre 2017 avec la polémique et les critiques de chercheurs<sup>28</sup> vis-vis de la validité des dernières évaluations à l'entrée au CP.

---

<sup>27</sup> Cette modélisation étant elle-même source potentielle de difficultés, comme nous l'avons vu au §5, avec les risques de confusion entre syllabes orales et syllabes écrites.

<sup>28</sup> Voir par exemple, sur le site Le Café pédagogique, les critiques d'E. Gentaz et L. Sprenger Charolles, « Evaluation de CP : un outil invalide », et celles de R. Brissiau. Repérées respectivement à : <http://www.cafepedagogique.net/Documents/14092017Article636409731527018464.htm> et à <http://www.cafepedagogique.net/Documents/13092017Article636408855185516203.htm>

## 7. Références

- 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.
- Baillé, J., & Maury, S. (Éds.) (1993). Les représentations graphiques dans l'enseignement et la formation. *Les sciences de l'éducation pour l'ère nouvelle, 93* (1-3) (numéro spécial).
- Bain, D (2002). De l'évaluation aux compétences : mise en perspective de pratiques émergentes. In J. Dolz & E. Ollagnier (Dir.), *L'énigme de la compétence en éducation* (pp. 129-145). Louvain-la-Neuve : De Boeck.
- Blanc, C. (2017). *Étude de la validité de l'évaluation interne conduite en Cours Préparatoire, en Français et en Mathématiques* (thèse de doctorat non publiée). Université Paris Descartes : Paris, France.
- Blanc, N., & Barrouillet, D. (2003). *Mémoire et compréhension – Lire pour comprendre*. Paris : In Press.
- Brousseau, G. (1988). Le contrat didactique : le milieu. *Recherches en Didactique des Mathématiques, 9*(3), 309-336.
- Carney, R.N., & Levin, J.R. (2002). Pictorial illustrations still improve students learning from text. *Educational Psychology Review, 14*(1), 5-26.
- Denis, M. (1989). *Image et cognition*. Paris : PUF.
- Devidal, M., Fayol, M., & Barrouillet, P. (1997). Stratégies de lecture et résolution de problèmes arithmétiques. *L'année psychologique, 97*(1), 9-31.
- 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.
- Elia, I. (2009). L'utilisation d'images dans la résolution de problèmes additifs : Quel type d'images et quel rôle ? *Annales de didactiques et de sciences cognitives, 14*, 5-29.
- Elia, I., Gagatsis, A., & Demetriou, A. (2007). The effects of different modes of representation 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 et A.B. Fuglestad (Eds.), *Proceedings of the 28th Conference of the International Group for the psychology of mathematics Education, 2* (pp. 327-334). Bergen, Norway : University College.
- Fijalkow, E. (2003). *L'enseignement de la lecture-écriture au cours préparatoire, entre tradition et innovation*. Paris : L'Harmattan.
- Goigoux, R. & Cèbe, S. (2013). Enseigner la compréhension à l'école élémentaire : des résultats de recherches à la conception d'un outil didactique. *Recherches, 58*, 29-46.
- Goigoux, R., Cèbe, S. & Paour, (2004). PHONO - Développer les compétences phonologiques - GS-CP. Paris : Hatier.
- Gonthier-Maurin, B. (2012). Sénat – *Rapport fait par la commission d'information sur le métier d'enseignant* (publication n°601, session ordinaire de 2011-2012).
- Grapin, N. (2015). Etude de la validité de dispositifs d'évaluation et conception d'un modèle d'analyse multidimensionnel des connaissances numériques des élèves de fin d'école (Thèse de doctorat non publiée). Université Paris-Diderot, Paris, France.
- Gyselinck, V. (1996). Illustrations et modèles mentaux dans la compréhension de textes. *L'année psychologique, 96*(3). 495-516.
- Jimenez, L., & Verschaffel, L. (2014). Development of children's solutions of non-standard arithmetic word problem solving. *Revista de Psicodidactica, 19*(1), 93-123.
- Johnson-Laird, P.N. (1983). *Mental Models*. Cambridge : Cambridge University Press.



- Johnson-Laird, P.N. (1993). La théorie des modèles mentaux. In Erlich, M.F., Tardieu, H. & Cavazza (Édit). *Les modèles mentaux : approche cognitive des représentations*. Paris : Masson.
- Julo, J. (1995). Représentation des problèmes et réussite en mathématiques. Rennes : PUR.
- Keenan, J. M., & Betjemann, R. S. (2006). Comprehending the Gray Oral Reading Test without Reading It: Why Comprehension Tests Should Not Include Passage-Independent Items. *Scientific Studies of Reading*, 10(4), 363-380.
- Kintsch, W. (1979). On modeling comprehension. *Educational Psychologist*, 14, 3-14.
- Leplat, J. (1997). Regards sur l'activité en situation de travail. Contribution à la psychologie ergonomique. Paris : PUF.
- Levin, J.R., Anglin, G.J. & Carney, R.N. (1987). On empirically validating functions of pictures in prose. In: Willows, D.M. & Houghton, H.A. (Eds.), *The Psychology of Illustrations : Basics Research* (Vol. 1, pp. 51-86). New York : Springer Verlag.
- Maury, S. (2002). A look at some studies on learning and processing graphic information, based on Bertin's theory. In Hitt, F. (Ed.) *Representations and Mathematics Visualisation. North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 297-310). Mexico: CINVESTA-IPN.
- MENJVA-MESR. (2011). *Note de synthèse sur le suivi de la réforme de l'enseignement primaire : pilotage local et suivi des élèves*.
- Monnier, N. (2003). Les schémas dans les activités de résolution de problèmes. *Grand N*, 71, 25-47.
- Nathan, M., Kintsch, W., & Young, E. (1992). A Theory of Algebra-Word-Problem Comprehension and Its Implications for the Design of Learning Environments. *Cognition and Instruction*, 9(4), 329-389.
- Peraya, D. (1995). Vers une théorie des paratextes : images mentales et images matérielles. *Recherches en communication*, 4, 1-38.
- Perrenoud, P. (2011). *Quand l'école prétend préparer à la vie... Développer des compétences ou enseigner d'autres savoirs ?* ESF : Issy-Les-Moulineaux.
- Reuter, T, Schnotz, W., & Rasch, R. (2015). Drawings and tables as cognitive tools for solving non-routine word problems in primary school. *American Journal of Educational Research*. 3(11), 1387-1397.
- Richard, J.-F. (2004, 4<sup>e</sup> éd.). *Les activités mentales – De l'interprétation de l'information à l'action*. Paris : Armand Colin.
- Riley, M.S., Greeno, J.G., & Heller, J.I. (1983) Development of children's problem solving ability in arithmetic. In H.P. Ginsburg (Ed.), *The Development of Mathematical Thinking*. New York : Academic Press.
- 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.
- Roditi, E., & Salles, F. (2015). Nouvelles analyses de l'enquête PISA 2012 en mathématiques : Un autre regard sur les résultats. *Éducation et formations*, 86-87, 236-267. Ministère de l'éducation nationale.
- Rudat, R., & Maury, S. (2009). Compréhension d'un graphique à double implantation par des élèves de 9 à 10 ans. *Aster*, 48, 187-208.
- Scallon, G. (2004). *L'évaluation des apprentissages dans une approche par compétences*. Bruxelles : De Boeck.
- Thevenot, C., & Coquin, D., & Verschaffel, L. (2006). La résolution de problèmes. In P. Barrouillet & V. Camos (Éds.), *La cognition mathématique chez l'enfant* (pp. 155-180). Paris : Solal.

- Tourneur, Y. (1979). A propos de l'appropriation des connaissances. In L. Allal, J. Cardinet & P. Perrenoud (Éds.), *L'évaluation formative dans un enseignement différencié* (pp. 182-186). Berne : Peter Lang.
- Vantourout, M., & Maury, S. (2017). Évaluation de la lecture au CP : mise en œuvre d'une approche multidimensionnelle. *Éducation et didactique*, 11(1), 45-62.
- Vantourout, M., & Goasdoué, R. (2014). Approches et validité psycho-didactiques des évaluations. *Éducation et Formation – e-302*, 139-156.
- Vergnaud, G. (1981). *L'enfant, la mathématique et la réalité*. Berne : Peter Lang.
- Vézin, L. (1986). Les illustrations, leur rôle dans l'apprentissage de textes. *Enfance*, 1, 109-126.
- Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation and autonomy in mathematics. *Journal for Research in Mathematics Education*, 27, 4, 458-477.