

NATHALIE SAYAC ET NADINE GRAPIN

EVALUER PAR QCM EN FIN D'ECOLE : STRATEGIES ET DEGRES DE CERTITUDE

Abstract. Multiple Choice Testing at the End of Primary School: Strategies and Levels of Confidence. The use of multiple choice items, especially in large-scale assessments, increasingly grows without questioning the strategies used by pupils to answer these types of questions, particularly in mathematics. The research presented in this paper aims to investigate how students in their final year of primary school proceed to answer multiple choice questions in the number domain and the level of confidence they give to their answer. Comparisons with students' levels of knowledge and used response strategies complement broader concerns.

Résumé. L'utilisation des QCM, en particulier dans les évaluations à grande échelle, se développe de plus en plus sans que soient interrogées les stratégies utilisées par les élèves pour répondre à ce type de questions, notamment en mathématiques. La recherche que nous présentons dans cet article vise à étudier la façon dont des élèves de fin de CM2 procèdent pour répondre à des QCM dans le domaine numérique et le degré de certitude qu'ils accordent à leur réponse. Des comparaisons selon le niveau de connaissance de l'élève et selon les stratégies de réponse qu'il emploie viennent compléter des considérations plus globales.

Mots-clés. QCM, évaluation, degré de certitude, stratégies.

Introduction

La culture scolaire française intègre progressivement les QCM (Questionnaire à choix multiple) comme mode d'évaluation alors que les anglo-saxons en ont une pratique beaucoup plus courante et ancienne. En mathématiques, dans le secondaire, ce type de support est désormais régulièrement utilisé pour évaluer les connaissances des élèves : par exemple, de nombreux manuels de mathématiques de collège intègrent les QCM comme modalités d'évaluation avec une visée diagnostique (en début de chapitre) ou comme auto-évaluation en fin de séquence. Parallèlement, si on considère les énoncés de l'épreuve de mathématiques du Diplôme National Brevet en 2012, sur les douze sujets recensés sur le site de l'APMEP, 7 comptaient des QCM. De plus, pour l'évaluation PISA 2003 (dominante mathématique), 20 % des questions étaient sous la forme de QCM simple (choisir une réponse parmi 4 propositions) et 12,9 % sous la forme de QCM complexe (succession de questions auxquelles il faut répondre par oui ou par non), c'est à dire, environ un tiers des questions de cette évaluation était sous forme de QCM (Bourny et al., 2003).

ANNALES de DIDACTIQUE et de SCIENCES COGNITIVES, volume 19, p. 169 - 197.
© 2014, IREM de STRASBOURG.

Ces premiers constats témoignent de l'intégration des QCM dans les activités mathématiques proposées aux élèves, au collège. Par contre, les élèves de l'école primaire ne sont que très rarement soumis à ce type de questions ou seulement lors d'activités exceptionnelles (rallyes, tests). Néanmoins, il convient de noter que pendant de nombreuses années le format QCM a été régulièrement employé dans les évaluations nationales qui avaient lieu à l'entrée en CE2 ou en 6^{ème} et plus récemment, en fin de CE1 ou en fin de CM2).

Ce format de questions est également largement répandu dans les évaluations menées à grande échelle par la DEPP (Direction de l'évaluation, de la prospective et de la performance), notamment dans le cadre des évaluations bilans de fin d'école et de collège relevant du CEDRE (Cycle des évaluations disciplinaires réalisées sur échantillon).

Alors que les résultats de ces évaluations nationales sont utilisés pour le pilotage pédagogique au niveau national et donnent lieu à la publication de notes d'informations rendant compte du niveau de connaissances et de compétences des élèves français dans les disciplines classiques, peu de travaux de recherche en France questionnent le dispositif d'évaluation par QCM pour ce qu'il est et pour ses spécificités, même si Adda (1976), Duval et Pluvinage (1977) et Pluvinage (1979), avaient commencé à s'intéresser à ces questions. Du côté anglo-saxon, Bloom, Hastings et Madaus (1971) ont été les précurseurs des questions posées par l'évaluation des apprentissages et ont été suivis par de nombreux chercheurs qui ont continué d'explorer ce champ. Néanmoins, les travaux de recherche qui portent sur les QCM concernent majoritairement des étudiants du supérieur (notamment Leclercq (1975, 1986), De Landsheere (1979), Hunt (1993), Gilles (1996)) alors que notre recherche porte sur des élèves d'un niveau scolaire très inférieur et d'âges beaucoup moins élevés (10-11 ans).

Ayant toutes deux participé à la conception et à l'analyse des évaluations bilans CEDRE 2008, nous nous sommes interrogées sur l'activité de l'élève de fin d'école lorsqu'il est confronté à des QCM. En effet, il nous a semblé pertinent de mener une recherche portant sur cette question à la fois par curiosité scientifique, mais aussi pour interpréter, avec plus de précision et de rigueur les résultats produits par ce type d'enquête. Dans cet article nous présentons la méthodologie que nous avons mise en œuvre et proposons, à titre expérimental, une utilisation de degrés de certitude dans le cadre d'une évaluation externe et sommative en fin de CM2 ; nous concluons par quelques constats en termes de fonctionnement cognitif des élèves et de pratique d'évaluation par QCM à l'école.

1. Problématique et hypothèses

L'utilisation de QCM dans des évaluations à grande échelle est largement répandue puisqu'elle permet une correction automatique (sans avoir recours à des personnes)

et fiable (pas d'interprétation des productions de la part de correcteurs par exemple).

En analysant l'ensemble des items proposés dans l'évaluation bilan CEDRE 2008, il s'est avéré que certains parmi ceux qui étaient proposés sous forme de QCM avaient des résultats très différents de ceux proposés sous forme de question ouverte, alors que d'un point de vue mathématique, la tâche à réaliser était équivalente. Il semblerait que, dans le cas d'un QCM, l'évalué est moins réticent à fournir une réponse que lorsque la question est posée sous forme ouverte : par exemple, pour l'évaluation PISA 2003, en France, les scores moyens de non-réponses étaient de 3,5% pour un QCM alors qu'ils étaient autour de 25,5 % pour une réponse ouverte construite (Bourny & al., 2003). Nous nous sommes donc demandé quelles pouvaient en être les raisons, ce qui nous a amenées à investiguer la façon dont les élèves appréhendaient ce dispositif spécifique d'évaluation. Nous avons souhaité observer les stratégies utilisées par les élèves de ce niveau de scolarité face à un QCM.

1.1 Stratégies de réponse

Tout d'abord, nous avons cherché à savoir si les stratégies que développent les élèves pour répondre aux QCM étaient toujours les mêmes et si elles étaient en lien avec leur niveau de connaissances en mathématiques. Est-ce que les élèves performants ou en difficultés adoptent une stratégie de réponse particulière aux QCM ? On trouve dans les travaux relatifs aux QCM (Choppin 1975, Leclercq 1987), différentes stratégies, mais toutes pensées pour des étudiants adultes. Voici, par exemple, les 3 modèles retenus par Choppin (1975) :

Modèle 1 : quand l'étudiant « sait », il choisit la réponse correcte et quand il ne « sait pas », il choisit au hasard parmi les réponses proposées.

Modèle 2 : commence comme le Modèle 1, mais au lieu de répondre au hasard quand il « ne sait pas », l'étudiant commence par éliminer les solutions qu'il sait être fausses et choisit au hasard parmi celles qui restent.

Modèle 3 : l'étudiant commence par ranger les solutions possibles par ordre de plausibilité décroissante et, si la consigne l'oblige à ne fournir qu'une d'entre elles, alors il choisit celle dont la probabilité (subjective) est la plus élevée (à ses yeux).

En complément de ces modèles, d'autres catégorisations existent, comme celle de Katz & al. (2000), qui distinguent des stratégies « *traditional* » et « *non traditional* ». Les premières reposent sur des procédures qui sont enseignées et que l'élève utiliserait pour répondre si la question était posée sous forme ouverte ; les autres, spécifiquement associées à des QCM, sont celles où l'élève s'appuie sur les réponses proposées pour faire un choix.

1.1.1 Liste des stratégies retenues

Nous avons donc commencé par lister les stratégies possibles pour des élèves de dix - onze ans car il nous a semblé indéniable qu'elles ne pouvaient être identiques à celles d'étudiants adultes. Cette liste s'est complétée au fur et à mesure de notre expérimentation, au plus près des stratégies réelles des élèves, pour arriver à la liste des stratégies suivante :

S1 : l'élève effectue la tâche demandée mentalement ou explicitement puis trouve, parmi les propositions, celle qui correspond à la réponse trouvée ;

S2 : l'élève reconnaît d'emblée la « bonne » réponse parmi celles proposées (connaissance intériorisée) ;

S3 : l'élève commence par s'engager dans une procédure de résolution, mais sans aller jusqu'au bout (à la différence de S1); il utilise ensuite les différentes propositions de réponses pour conclure (en choisissant celle qui lui paraît la plus vraisemblable) ;

S4 : l'élève élimine les propositions qui paraissent fausses, puis déduit de celle(s) qui reste(nt), la bonne réponse ;

S5 : l'élève répond au hasard ;

S6 : l'élève passe en revue superficiellement toutes les propositions, puis choisit celle qui lui paraît la plus vraisemblable ;

S7 : l'élève ne sait pas expliquer sa procédure ;

S8 : l'élève combine les nombres en présence de manière à trouver, parmi les choix possibles, une solution ;

S9 : l'élève applique une règle simple intériorisée, correcte ou non (théorème en actes) ;

S10 : l'élève teste les propositions de réponse une à une jusqu'à trouver celle qui convient.

1.1.2 Groupes de stratégies

Afin d'analyser plus efficacement le lien entre les différentes stratégies des élèves en fonction de leur niveau de connaissance, nous avons effectué des groupements de stratégies par type :

Stratégies A (S1, S2, S9, S10) : stratégies de savoirs. L'élève active des connaissances ou des savoir-faire (techniques – raisonnement) pour choisir la réponse qu'il pense être la bonne : soit il résout complètement la tâche (par la procédure de son choix, juste ou fausse), soit il teste les propositions de réponse et choisit celle qui peut convenir. Les stratégies S1, S2 et S9 peuvent s'apparenter aux

« *traditional* » stratégies définies par Katz & al (2000), alors que la stratégie S10 est spécifique au format QCM.

Stratégies B (S5, S6, S7, S8) : stratégies de substitution ou de repli. L'élève n'utilise pas ses connaissances mathématiques de façon explicite pour faire un choix : son choix ne repose pas de façon assurée sur ses connaissances.

Stratégies C (S3, S4) : stratégies mixtes. L'élève a initié un raisonnement pour répondre à la question posée, mais il se sert des différentes propositions de réponse pour finaliser son choix.

1.2 Degrés de certitude et de réalisme

En complément des stratégies utilisées, des procédures mises en œuvre et de la réponse choisie, il nous a semblé complémentaire de pouvoir demander à l'élève d'estimer la qualité de sa réponse (bonne ou non) en lui demandant quel degré de certitude il lui attribuait. Définis par Leclercq et Poumay (2004, p. 3), les degrés de certitude sont « l'ensemble des jugements, des analyses, des régulations, conscientes ou non, (mais qu'il importe de rendre explicites, observables et conscientes) effectués par l'apprenant sur ses propres performances ».

En nous inspirant des travaux de Leclercq (1987, 2006), nous avons proposé une échelle de quatre degrés de certitude, adaptée¹ à des élèves de fin d'école et variant de « pas sûr du tout » à « sûr et certain » : pas sûr du tout (1), pas très sûr (2), sûr (3), sûr et certain (4). En effet, même si des travaux de recherche antérieurs ont stipulé que la consigne ne devait pas être verbale, mais devait plutôt être présentée sous forme de pourcentage (0%, 20 %...) car la subjectivité liée aux expressions verbales était problématique (Gilles, 1996 et Leclercq dès 1983), nous n'avons pas suivi cette préconisation. Nous avons en effet considéré que, pour des élèves de cet âge qui ne maîtrisent pas les pourcentages, il était plus opportun de leur présenter une consigne verbale à quatre degrés de certitude, accompagnée d'une échelle numérique (de 1 à 4), en leur laissant la possibilité de positionner leur degré de certitude entre deux degrés « entiers » (c'est-à-dire choisir un degré de certitude de 3,5 par exemple). La présentation de l'échelle choisie (voir Annexe 1) permettait ainsi de doubler la consigne verbale d'une consigne numérique, tout en évitant le recours aux pourcentages. Même s'il est « illusoire » d'avoir une précision supérieure à 20 %, (Leclercq, 2006), c'est à dire correspondant à six degrés de certitude entre 0 et 100 %, l'échelle que nous avons construite conduit à sept degrés possible si l'on permet de faire des croix entre deux des quatre degrés indiqués. L'élève, après avoir choisi une réponse pour chaque item, devait donc estimer la qualité de celle-ci à l'aide de ces degrés réaménagés.

¹ Les degrés de certitude proposés par Leclercq varient de 0% à 100%, par tranches de 20%.

Ainsi, en mettant en perspective la réponse choisie (juste ou fausse) et le degré de certitude que l'élève lui confère, nous pouvons savoir si l'élève est réaliste ou non quant à la qualité de sa réponse. Boulé & Laveault (2011) qui utilisent les degrés de certitude et de réalisme dans une évaluation diagnostique expliquent que « l'étudiant réaliste tend à anticiper une bonne réponse lorsque celle-ci est bonne et une réponse incorrecte lorsque celle-ci est incorrecte ». Les recherches menées par Leclercq (1993) et citées par Boulé & Laveault (2011) laissent penser que l'utilisation de degrés de certitude et de réalisme avec des étudiants peut avoir une fonction formative : « l'indice de réalisme obtenu fournit un feedback à l'étudiant et lui permet d'améliorer son estimation de ses compétences ». Nous n'avons pas utilisé les degrés de certitude à cette fin puisqu'ils accompagnaient une évaluation externe, mais plutôt pour estimer l'indice de réalisme dont les élèves de fin d'école pouvaient faire preuve.

Nous avons également estimé que l'adjonction du degré de certitude à une réponse pouvait nous donner des informations sur la connaissance (ou la méconnaissance) d'un élève sur un savoir donné. En effet, pour chaque item la nature de la question posée et des distracteurs² est didactiquement intéressante pour repérer ce qui est vraiment en jeu du point de vue des connaissances. Par exemple, relativement au domaine des nombres décimaux, les élèves ayant une représentation erronée des décimaux et qui associent $62/10$ à $62,10$ avec un degré de certitude supérieur à 3 auront plus de difficulté à déconstruire cette (mé)connaissance que les élèves qui choisissent $62,10$ en étant peu sûrs (degré de certitude compris entre 1 et 2).

A la différence des travaux menés dans le cadre de la docimologie sur les degrés de certitude et de réalisme, comme la recherche évoquée précédemment, nous n'avons pas questionné globalement la cohérence du test ou la fiabilité de nos items, mais, de façon complémentaire, sur un nombre restreint d'items, nous avons cherché à observer comment s'articulaient les degrés de certitude exprimés par les élèves et leur niveau de connaissances en mathématiques ou les stratégies qu'ils adoptent pour répondre à un QCM. En France, à ce niveau scolaire (fin d'école primaire, élèves âgés de 10-11 ans), peu d'études, à notre connaissance, ont été menées pour savoir comment les élèves traitent ce type de question.

Cette recherche permet également d'étudier les « risques » (Pluvinage, 1979) liés à ce type de format de question, ce qui n'est pas sans intérêt. En effet, dans le cadre d'une évaluation institutionnelle de fin de niveau scolaire, il est indispensable de minimiser les risques d'attribuer un résultat négatif à un élève qui a acquis la connaissance ou les compétences en jeu et réciproquement car les résultats des

² Dans un QCM, un distracteur est une des réponses proposées qui n'est pas la bonne.

bilans tels que ceux produits dans CEDRE³ peuvent contribuer aux décisions qui orientent les politiques éducatives.

Après avoir décrit la méthodologie employée, nous présentons en premier des résultats globaux portant sur les stratégies employées par les élèves pour répondre aux différents items, selon leur niveau de connaissance et en lien avec le degré de certitude qu'ils accordent à leurs réponses. Une deuxième partie porte sur les stratégies employées, les degrés de certitude accordés aux réponses selon la qualité de la réponse fournie (degré de réalisme) et le niveau de connaissance de l'élève. Logiquement, nous faisons l'hypothèse que les stratégies de savoir sont plus employées par les élèves « forts », qu'elles conduisent davantage à la bonne réponse que les autres stratégies et qu'elles sont accompagnées de degrés de certitude élevés. Cette hypothèse est-elle valable sur l'ensemble des items choisis ? Quelles sont les stratégies mises en œuvre par les élèves moyens » ou « faibles » ? Avec quel degré de certitude fournissent-ils leur réponse ?

Enfin, ces résultats généraux seront complétés dans une troisième partie par une analyse plus locale sur les écritures décimales ; elle nous permettra d'interroger les apprentissages qui peuvent être (ou non) évalués par QCM et montrera l'intérêt de l'utilisation des degrés de certitude et de réalisme dans ce type de résultats.

2. Expérimentation

2.1 Principe

L'expérimentation que nous avons menée s'est déroulée en juin 2012, dans six classes de CM2 de Paris et de Montreuil. Elle a eu lieu après les évaluations nationales en mathématiques (début juin) afin de pouvoir bénéficier des résultats des élèves et ainsi déterminer leur niveau de connaissances de manière objective et permettant des comparaisons. Nous avons convenu d'élaborer un test comportant un nombre limité d'items afin que chaque élève puisse nous expliciter son raisonnement individuellement lors de la passation, et ainsi être en capacité de déterminer au plus juste la stratégie employée par l'élève pour choisir sa réponse.

2.2 Classes

Nous avons souhaité réaliser notre expérimentation dans des classes variées en termes de population et de niveau scolaire afin de pouvoir tester toutes nos hypothèses. Ainsi, nous avons retenu 2 classes des « beaux » quartiers parisiens,

³La direction de l'évaluation, de la prospective et de la performance qui conçoit les bilans CEDRE exerce une fonction de suivi statistique, d'expertise et d'assistance pour le Ministère de l'éducation nationale et le ministère de l'enseignement supérieur et de la recherche.

deux classes « banales » (1 à Paris, 1 à Montreuil), 2 classes ZEP ou « éclair » (1 à Paris, 1 à Montreuil), soit 155 élèves au total.

2.3 Test

Le test a été conçu à partir de sept items (énoncés figurant en Annexe 1) ; trois d'entre eux (items 1, 4 et 6) sont extraits de l'évaluation bilan CEDRE 2008 pour lesquels les scores de réussite se distinguaient des résultats habituels en mathématiques. Pour la formulation des questions et le choix des distracteurs⁴ nous nous sommes appuyées sur les « règles » déterminées dans différents ouvrages, en particulier ceux de Leclercq (1986) et Grégoire & Laveault (1997) ; nous avons donc principalement retenu des distracteurs « vraisemblables » afin qu'ils ne soient pas éliminés immédiatement et qu'ils correspondent à des erreurs fréquentes d'élèves. En effet, si les distracteurs ne correspondent pas à des erreurs classiques, le format QCM sera plus simple que la question ouverte, alors que dans le cas contraire, les deux formats pourraient être équivalents du point de vue de la difficulté (Katz, et al., 2000), sous certaines conditions. Pour les items relevant des fractions-décimaux, les différents choix de réponses ont été effectués à partir d'erreurs reconnues comme fréquentes par différents travaux en didactique des mathématiques (Douady & Perrin-Glorian (1986), Comiti et Neyert (1979), Roditi (2007)).

En choisissant des distracteurs correspondant à des erreurs fréquentes d'élèves, nous souhaitons également savoir s'ils étaient attractifs pour certains élèves (selon leur niveau de connaissances) ou au contraire s'ils étaient d'emblée rejetés. L'utilisation des degrés de certitude, notamment pour les réponses correspondant à des conceptions erronées sur les fractions-décimaux nous permettra, au-delà du score de réussite, de préciser si l'élève est réaliste ou non du point de vue de ses connaissances. En Annexe 2 figure une analyse didactique du choix des distracteurs proposés⁵.

Il convient de noter que nous avons opté pour des QCM avec 4 choix de réponses possibles pour rester dans le format du bilan CEDRE.

Observer des élèves pendant qu'ils résolvaient ces problèmes spécifiques devait ainsi nous permettre de mieux comprendre les scores de réussite. Nous avons choisi les sept items pour qu'ils soient de niveaux variés du point de vue de la difficulté (l'item 7 étant particulièrement difficile) et pour qu'ils relèvent de deux catégories : ceux pour lesquels l'élève doit mobiliser des connaissances relatives

⁴Nous précisons que dans les items que nous proposons, comme dans ceux figurant dans l'évaluation CEDRE, chaque question possède une seule bonne réponse.

⁵ Cette analyse est menée *a posteriori* pour les items 1, 4 et 6 que nous avons extraits de l'évaluation ; nous n'avons pas choisi les distracteurs *a priori*.

aux fractions et aux nombres décimaux et ceux pour lesquels l'élève doit résoudre un problème dans le cadre des grandeurs et mesures (problèmes additifs, de proportionnalité et de calcul de périmètre).

Un espace libre a été prévu pour chaque item afin que les élèves puissent éventuellement faire des calculs.

2.4 Passation

Une présentation à la classe entière a été proposée avant la passation individuelle pour expliquer la nature du test et le déroulement de l'entretien, en précisant les consignes : « pour chaque exercice, il faut trouver la bonne réponse parmi celles proposées et indiquer son degré de certitude ». Nous avons choisi de faire des entretiens individuels pour que les élèves puissent nous expliquer précisément comment ils avaient choisi leurs réponses et leurs degrés de certitude. Nous avons fait passer le test aux élèves un par un⁶ et noté toutes les explications ou remarques qu'ils avaient formulées.

A la fin du test, nous avons demandé aux élèves s'ils avaient déjà fait des QCM et ce qu'ils pensaient de ce type d'évaluation.

2.5 Niveaux de connaissance des élèves

Enfin, pour déterminer les niveaux de connaissance des élèves, nous avons utilisé leurs résultats aux évaluations nationales de mathématiques qu'ils venaient tous de passer quelques jours avant notre test. Nous avons ensuite classé les élèves suivant trois groupes en fonction de leurs résultats à ces évaluations :

- Élèves « faibles » : ceux qui ont moins de 40% de réussite (32 élèves),
- Élèves « moyens » : entre 40% et 75% de réussite (64 élèves),
- Élèves « forts » : ceux qui ont plus de 75% de réussite (59 élèves).

3. Résultats

3.1 Résultats globaux

Les scores de réussite par item figurent dans le tableau 1, ci-dessous. Le premier constat que nous pouvons faire concerne l'aspect spécifique du fonctionnement cognitif des élèves en situation. Il s'avère en effet qu'il est parfois extrêmement difficile d'appréhender leur fonctionnement cognitif car il ne s'inscrit pas dans la même logique que celui d'un adulte. Il n'a d'ailleurs pas toujours été facile de

⁶ Nous nous sommes placées dans un espace proche de la classe (couloir ou salle à côté) afin que les élèves effectuent leur test, sans perturber le déroulement prévu de la classe.

comprendre comment un élève était arrivé à faire tel ou tel choix car il n'était pas capable de l'expliquer lui-même.

Pour formaliser les résultats de l'évaluation-bilan CEDRE 2008, la DEPP a construit une échelle de performance⁷ à six niveaux selon la méthodologie décrite dans la Note d'information n°10-18 : « La partie la plus basse de l'échelle est constituée des scores obtenus par les 15% d'élèves ayant les résultats les plus faibles. À l'opposé, la partie supérieure, constituée des scores les plus élevés, rassemble 10% des élèves. Entre ces deux niveaux, l'échelle a été scindée en trois parties d'amplitude de scores égale correspondant à trois groupes intermédiaires». Les élèves sont donc répartis en 6 groupes : du groupe 0 (les plus faibles) au groupe 5 (les plus performants).

Les items 1, 4 et 6 extraits de l'évaluation CEDRE 2008 permettent de caractériser certaines connaissances des élèves d'un groupe donné : les items 1 et 4 correspondent à celles du groupe 3 et l'item 6 à celles du groupe 4. Ce qui signifie, d'après les résultats de l'évaluation-bilan CEDRE, que 58% des élèves de fin de CM2 ont une probabilité supérieure à 50% de réussir les items 1 et 4, alors que seulement 28% des élèves ont la même probabilité de réussir l'item 6. Les résultats que nous obtenons dans ce test pour les items 1, 4 et 6 sont ainsi globalement cohérents avec ceux de l'évaluation-bilan. En comparant les items 2, 3 et 5 à des items *a priori* équivalents du point de vue de la tâche et de la complexité et figurant dans l'évaluation bilan, on peut aussi constater que les scores de réussite obtenus dans ce test sont également cohérents avec ceux de l'évaluation bilan.

	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7
Score de réussite	58,7%	51,6%	54,8%	71,6%	50,3%	46,5%	27,1%

Tableau 1 - Score de réussite moyen pour chacun des items.

Les difficultés relevées sur les connaissances des nombres décimaux par de nombreuses recherches en didactique des mathématiques sont aussi confirmées à nouveau par les scores de réussite aux items 3, 4 et 6.

La plupart des élèves n'avaient jamais été confrontés à des QCM, si ce n'est en français ou en anglais comme certains nous l'ont indiqué. Dans l'ensemble, ce type de dispositif d'évaluation leur a plu. Certains ont estimé qu'il était plus facile que ce qu'ils avaient l'habitude de faire, d'autres plus difficile, mais ces considérations

⁷ Des précisions sur cette évaluation et sur le contenu de l'échelle figurent dans la Note d'information n°10-18 (éditée par le Ministère de l'éducation nationale) ou dans la brochure "Mathématiques en fin d'école primaire : le bilan des compétences", éditée par le Sceren.

se rapportaient essentiellement aux questions posées et non à la modalité d'évaluation, nous a-t-il semblé.

3.2 Stratégies employées par les élèves

3.2.1 Choisir « au hasard »

Précisons qu'il est important de distinguer, ce qui relève effectivement du hasard (*random-guessing* ou *blind-guessing*) de stratégies permettant de deviner la bonne réponse (*guessing*) : « si un choix aléatoire est aveugle, un choix deviné ne l'est pas » (Leclercq, 1987, p. 13). Par ailleurs, le choix « au hasard » n'est retenu que par les élèves qui ne savent pas répondre, donc généralement par les plus faibles. Or, il s'avère, pour les élèves les plus faibles, que s'ils répondaient au hasard (c'est à dire dans un QCM à quatre choix, avec une probabilité de 1/4 de trouver la bonne réponse), ils auraient de meilleurs résultats que ceux qu'ils obtiennent (Grégoire & Laveault, 1997). Les résultats que nous proposons ici dans le tableau 2 s'appuient sur le déclaratif des élèves quant à leur procédure « au hasard », et ne permettent pas de discerner véritablement ce qui relève du hasard et ce qui relève du « *guessing* ».

	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7
Réponse "au hasard"	1,9 %	1,3 %	8,4 %	9,7 %	3,9 %	9 %	12,9 %

Tableau 2 - Pourcentage d'élèves ayant déclaré répondre "au hasard" pour chacun des items.

Environ 2% des élèves déclarent choisir une réponse au hasard pour répondre aux items 1, 2 et 3, c'est à dire ceux correspondant à un problème peu complexe. Si on considère les autres items, environ 10% des élèves semblent répondre au hasard principalement dans deux cas : face à un problème complexe (item 7) pour lequel ils n'arrivent pas à mener à terme une procédure de résolution ou pour répondre à une question dont ils ignorent totalement quelle pourrait être la réponse (notamment items 3 et 4 portant sur les fractions /décimaux).

La recherche que nous avons menée tend à montrer que finalement peu d'élèves déclarent répondre « au hasard » dans ce type de dispositif. De façon globale, les élèves ne semblent pas avoir perçu les avantages que ce dispositif pouvait leur procurer lorsqu'ils ne connaissaient pas la réponse et ont été peu nombreux à utiliser la stratégie S5 (au hasard) ; ils ont volontiers expliqué qu'ils utilisaient d'autres stratégies de repli, correspondant davantage à du « *guessing* » qu'à du hasard. Il semblerait de plus qu'un effet de contrat didactique empêche les élèves de fin d'école primaire à considérer que « répondre au hasard » puisse être

envisageable pour répondre, même dans les rares cas où nous leur avons suggéré face à leur désarroi.

3.2.2 Stratégies utilisées par item et scores de réussite

Globalement, comme l'indique le tableau 3, pour répondre aux QCM, les élèves utilisent majoritairement des stratégies de savoir (type A). Dans le cas particulier de l'item 7 (résolution d'un problème complexe), les élèves procèdent différemment. Une analyse plus fine des stratégies utilisées pour cet item montre que 12% des élèves combinent les nombres en présence pour obtenir une des réponses proposées, ce qui n'est pas le cas face à un problème de moindre complexité (item 5) où seulement 1,3 % des élèves procèdent ainsi.

	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7
Stratégies A	73,5%	89,0%	68,4%	79,4%	63,9%	60,6%	41,9%
Stratégies B	16,1%	7,7%	26,5%	18,1%	8,4%	25,2%	43,9%
Stratégies C	10,3%	3,2%	5,2%	2,6%	27,7%	14,2%	14,2%

Tableau 3 - Répartition des stratégies utilisées par les élèves selon les items.

Pour les items mettant en jeu des connaissances sur les nombres décimaux (items 3, 4 et 6), les élèves appliquent majoritairement une règle simple intériorisée qui les mène ou non à la bonne réponse (stratégie S9) : dans ces cas, les élèves appliquent plus ou moins justement une technique (décalage de la virgule ou ajout de zéros...) afin de trouver une réponse. Par exemple, pour trouver le quotient de 872 par 100 (item 4), 58,7% des élèves utilisent une stratégie S9, mais parmi eux, seulement 77% trouvent la bonne réponse.

La spécificité du dispositif QCM influence particulièrement les stratégies des élèves lorsqu'ils sont face à un problème difficile à résoudre pour eux : la possibilité de choisir une réponse parmi d'autres modifie apparemment les procédures de résolution (*guessing* ou combinaison des nombres en jeu pour obtenir un des choix proposés) : ce qui a déjà été explicité pour des étudiants, notamment dans Katz & al (2000), et dans des études visant à comparer les effets du format de question (ouvert ou QCM) sur la réussite, telles que celles recensées dans Simkin & Kuechler (2005).

3.2.3 Changement de stratégie au cours du test

La nature des questions posées et des distracteurs proposés n'est évidemment pas sans incidence sur les stratégies que déploient les élèves pour y répondre ; le tableau 4 nous permet d'étudier la variété des stratégies adoptées par des élèves de fin d'école confrontés à des QCM.

	Stratégies A	Stratégies B	Stratégies C
7 fois	13,5%	0,0%	0,0%
6 fois	21,3%	0,0%	0,0%
1 fois	0,0%	27,7%	32,9%
0 fois	0,0%	29,7%	49,0%

Tableau 4 - Répartition des élèves selon le nombre de fois où ils ont utilisé une stratégie donnée sur l'ensemble du test.

Globalement, les élèves utilisent des stratégies différentes suivant les items auxquels ils sont confrontés, notamment selon leur niveau de connaissances (comme nous le verrons dans le paragraphe suivant). En effet, on peut relever que seulement 13,5% des élèves ne changent pas de type de stratégies sur l'ensemble des items et que ces élèves n'utilisent que des stratégies de type A. Par ailleurs, on peut également noter que 30% des élèves n'utilisent jamais de stratégies de type B, c'est à dire que pour répondre, ils utilisent forcément des connaissances ou a minima, initient un raisonnement.

Nous en concluons que les élèves de fin d'école savent tirer avantage du format QCM, même s'ils ne sont pas habitués à ce type de dispositif : ils n'hésitent pas à utiliser des stratégies de repli ou de substitution (en utilisant éventuellement les ordres de grandeurs) lorsqu'ils ne connaissent pas la réponse ou encore à se servir des réponses proposées pour les tester comme solutions potentielles (stratégie S10 utilisée par exemple dans l'item 2 pour trouver la longueur du rectangle en testant les différentes valeurs proposées).

3.2.4 Stratégies utilisées et niveaux de connaissance

Afin de considérer dans quelle mesure les stratégies utilisées par les élèves étaient conditionnées par leur niveau de connaissances, nous avons répertorié le type de stratégies utilisé par les élèves suivant leur groupe d'appartenance pour chaque item (Annexe 3) et pour l'ensemble du test (tableau 5 ci-dessous).

	Stratégies A	Stratégies B	Stratégies C
Élèves « faibles »	64,9%	25,4%	9,7%
Élèves « moyens »	63,5%	24,7%	11,8%
Élèves « forts »	76,1%	12,8%	11,1%

Tableau 5 – Répartition des stratégies utilisées par les élèves selon leur groupe d'appartenance pour l'ensemble du test

En moyenne, les élèves les plus faibles et ceux du groupe intermédiaire utilisent des types de stratégies assez similaires ; les élèves les plus forts utilisent deux fois moins de stratégies de repli ou de substitution que les autres élèves. C'est

seulement à partir d'un niveau de connaissance élevé (élèves « forts ») que les élèves utilisent prioritairement des stratégies qui relèvent de la connaissance, alors que les élèves « moyens » ou « faibles » emploient davantage de stratégies mixtes ou de repli.

En observant plus finement item par item (Annexe 3), on peut constater que, logiquement, les stratégies de type A sont davantage utilisées par les élèves ayant un plus haut niveau de connaissances (jusqu'à 98%), sauf pour les items 3 et 6 (lien fraction/décimaux) où les élèves les plus faibles sont les plus nombreux à utiliser ce type de stratégies, même si les techniques qu'ils appliquent sont erronées.

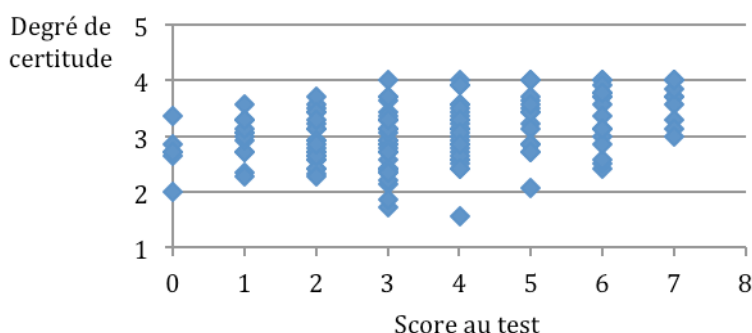
De la même manière, les élèves qui utilisent davantage des stratégies de substitution ou de repli se trouvent parmi les élèves les plus faibles (jusqu'à 61% pour l'item 7, le plus complexe du test), même si pour les items 3 et 6, ce sont les élèves de niveau « moyen » qui y ont le plus recours (36% et 33% contre 13% et 16% pour les plus faibles).

3.3 Degrés de certitude

Les élèves se sont très bien appropriés les degrés de certitude (de 1 à 4 suivant qu'ils étaient peu ou très sûrs d'eux) et n'ont pas hésité à faire une croix entre deux degrés quand cela leur semblait plus proche de ce qu'ils ressentaient. Il n'y a eu aucun refus à renseigner ces degrés et les hésitations étaient peu nombreuses.

3.3.1 Degrés de certitude et réussite

Le nuage de points ci-dessous représente le degré de certitude en fonction du score de réussite au test (nombre de réponses justes sur l'ensemble du test) ; le score varie entre 1 et 7.



Sur l'ensemble des sept items, la réussite globale au test est corrélée avec le degré de certitude ($r = 0,393$), ce qui signifie que plus l'élève a un score de réussite important, plus il est sûr de ses réponses. Ce résultat nous paraît assez logique,

mais nous avons souhaité l'affiner, notamment en étudiant plus spécifiquement les degrés de certitude selon les différents groupes d'élèves (forts, moyens et faibles) et les différents items.

3.3.2 Degrés de certitude et niveaux de connaissance des élèves

Pour chaque item, nous avons souhaité étudier comment les élèves répondaient et avec quel degré de certitude, en les distinguant suivant leur niveau de connaissances. Le tableau 6 ci-dessous indique les scores de réussite par items et les degrés de certitude attribués à chacun des items par les trois groupes d'élèves :

		Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Moyenne
Élèves faibles	SR	25,8 %	35,5 %	32,3 %	41,9 %	19,4 %	19,4 %	19,4 %	27,7%
	DC	2,74	3,13	3,06	2,74	3,11	3,06	2,35	2,887
Élèves moyens	SR	54 %	46%	58,7 %	73 %	39,7 %	44,4 %	19 %	47,8%
	DC	2,92	3,52	2,67	3,10	3,54	2,75	2,35	2,974
Élèves forts	SR	81 %	67,2 %	65,5 %	86,2 %	79,3 %	63,8 %	39,7 %	68,9%
	DC	3,30	3,70	3,03	3,61	3,65	3,04	2,90	3,319

Tableau 6 - Scores de réussite (SR) et degrés de certitude (DC) par item et par groupe d'élèves

Logiquement, les élèves les plus performants sont plus sûrs de leurs réponses que les élèves qui le sont moins (en moyenne, 2,887 pour les plus « faibles », 2,974 pour les « moyens », 3,319 pour les plus « forts »). Il convient de noter que le degré moyen de certitude est relativement élevé pour les trois groupes d'élèves, et que celui du groupe « forts » se démarque de façon significative⁸ de ceux des deux autres groupes qui sont eux, assez proches.

Pour les deux items 3 et 6 relatifs au lien fractions/décimaux : les élèves du groupe le plus faible ont le degré de certitude le plus élevé des trois groupes (3,06) alors que leur score moyen de réussite pour ces deux items n'est respectivement que de 32,3% et 19,4%. Ces élèves utilisent des théorèmes en actes faux dont ils sont

⁸ La comparaison des moyennes des DC entre les différents groupes a été réalisée en calculant le z de la loi normale. Pour la différence entre DC (élèves forts - élèves moyens), $z = 3,734$: ce qui est significatif ; alors que pour la différence entre DC (élèves moyens - élèves faibles), $z = 1,438$: ce qui n'est pas significatif.

certaines de leur validité (ils associent généralement la fraction a/b au nombre décimal a,b), ce qui est préoccupant. Nous reviendrons plus spécifiquement sur les items 3, 4 et 6 dans le paragraphe 3.4 portant sur les écritures décimales.

On peut être étonné aussi de relever que chez les élèves du groupe « faible », l'item 5 peu réussi (moins de 20%) soit aussi affecté d'un degré de certitude supérieur à 3. Comme pour les deux items précédents relatifs aux liens fractions/décimaux, cela signifie que ces élèves ne réalisent pas que leurs connaissances mathématiques ne sont pas suffisantes pour réussir les tâches proposées. Pour ces trois items, les élèves accordent un degré de certitude important à leur réponse parce qu'elle correspond à un des choix proposés parmi les distracteurs : ici, pour l'item 5, les élèves qui ont choisi 3,60 € (résultat intermédiaire de la résolution du problème) étaient sûrs de leur choix puisque le résultat trouvé figurait parmi la liste des réponses proposées.

Enfin, l'item 7, qui est le plus difficile, a logiquement des degrés de certitude les plus bas, pour les trois groupes.

Les recherches sur le sentiment d'efficacité personnelle (Bandura (2003), Rondier, (2004)) peuvent nous amener à penser que les élèves en difficulté n'ont pas confiance en eux et que cela impacte leur réussite à l'école, mais nous constatons à travers cette enquête, que ce n'est pas forcément ce sentiment qui est à l'origine de l'échec de ces élèves et qu'il faudrait relativiser cette croyance. L'utilisation des degrés de réalisme obtenus en croisant le degré de certitude à la qualité de la réponse de l'élève peut permettre de questionner différemment les difficultés des élèves pour mieux les prendre en compte dans l'enseignement (ce que nous ferons localement dans le paragraphe 3.4 pour les items de ce test portant sur les fractions et les nombres décimaux).

3.3.3 Stratégies et degrés de certitude

Mettre en perspective le degré de certitude accordé à une réponse avec la stratégie employée pour la trouver nous semble intéressant pour éclairer le rapport au savoir mathématique que les élèves de fin de primaire peuvent avoir : est-ce qu'ils sont plus sûrs de leur réponse lorsqu'ils emploient une stratégie de savoir (A) que lorsqu'ils emploient une stratégie d'évitement (B) ? Est-ce toujours le cas ?

Nous avons synthétisé dans le tableau 7, relativement aux stratégies employées, les degrés de certitude qui étaient accordés à la réponse. Ainsi, pour les élèves qui ont employé une stratégie A pour répondre à l'item 1, 13 % d'entre eux accordent un DC inférieur à 2 et 77 % accordent un DC supérieur à 3 (le reste accordant un DC égal à 2,5).

	Stratégie A		Stratégie B		Stratégie C	
	$0 \leq DC \leq 2$	$3 \leq DC \leq 4$	$0 \leq DC \leq 2$	$3 \leq DC \leq 4$	$0 \leq DC \leq 2$	$3 \leq DC \leq 4$
Item 1	13,2%	77,2%	76,0%	4,0%	43,8%	18,8%
Item 2	5,8%	89,1%	25,0%	58,3%	20,0%	60,0%
Item 3	20,7%	74,5%	58,5%	17,1%	62,5%	25,0%
Item 4	8,1%	86,9%	71,4%	14,3%	100,0%	0,0%
Item 5	6,1%	90,9%	69,2%	30,8%	13,9%	83,7%
Item 6	18,1%	72,3%	64,1%	30,8%	40,9%	50,0%
Item 7	27,7%	58,5%	58,8%	26,5%	54,5%	31,8%

Tableau 7 – Degrés de certitude (DC) relativement aux stratégies employées.

De façon globale, pour tous les items, les élèves qui emploient une stratégie A accordent un degré de certitude élevé à leur réponse (mis à part pour l'item 7) ; en revanche, les élèves qui emploient des stratégies de repli (B) ou mixtes (C) n'accordent pas de façon aussi systématique un faible degré de certitude à leur réponse. En effet, principalement pour l'item 5 qui correspond à la résolution d'un problème, les élèves qui emploient une stratégie mixte (C) sont 83,7 % à être sûrs de leur réponse.

Au vu de ces quelques résultats, et avec toutes les précautions qu'il faut prendre relativement à la taille à notre échantillon, il semblerait, notamment en situation de résolution de problèmes sous forme QCM, que les élèves de fin d'école sont en capacité de trouver dans d'autres stratégies que celles de savoir une certaine efficacité pour répondre. Ce type de stratégies, spécifiquement lié aux QCM, ne peut certainement pas fonctionner dans le cas d'une question ouverte.

3.4 Évaluation des connaissances des élèves sur les nombres décimaux

Au-delà des résultats globaux que nous avons présentés précédemment, nous souhaitons détailler dans ce paragraphe ce que les trois items relatifs aux nombres décimaux nous ont permis de mettre en évidence car les résultats attachés à ces items sont à la fois surprenants et instructifs et nous permettent d'étudier plus finement le comportement des élèves de fin d'école confrontés à des QCM.

3.4.2 Passage d'une écriture fractionnaire à une écriture à virgule et réciproquement : réussite et échec

Les items 3 et 6 du test permettent de révéler les apprentissages des élèves sur l'écriture d'un nombre décimal et les difficultés qu'ils peuvent rencontrer dans deux tâches différentes : transformer une écriture sous forme de fraction décimale en une écriture à virgule (item 3) et réciproquement (item 6).

Un tableau croisé des réponses à chacun de ces deux items (Annexe 4, tableau 1) nous conduit à un premier constat : seuls 38 % des élèves réussissent cette transformation dans les deux sens, ce qui pourrait être considéré comme un indice de connaissance maîtrisée (alors que les items 3 et 6 ont respectivement un score de réussite de 54,8 % et 46,5 %). Ce constat met en évidence que d'une part, la connaissance des décimaux n'est pas achevée en fin d'école, et que d'autre part les scores de réussite à ces deux items sont à interpréter avec prudence et réserve.

Par ailleurs, dans chacun de ces deux items, un des distracteurs correspond à l'erreur signalée précédemment consistant à considérer le trait de fraction et la virgule comme deux séparateurs similaires (c'est-à-dire à concevoir $237/8$ comme $237,8$ et $62,10$ comme $62/10$) : nous constatons que 18 % des élèves ont cette conception sur les deux items, soit presque un élève sur cinq ! Même si à la fin du cycle 3, les apprentissages relatifs à l'écriture des nombres décimaux sont encore en cours de construction et que l'utilisation des QCM à ce stade de l'apprentissage peut surévaluer les connaissances de l'élève (Grégoire & Laveault, 1997), il n'empêche que ces résultats sont révélateurs d'une difficulté qui perdure.

Nous voyons ici l'intérêt de croiser les réponses à ces deux items pour préciser les connaissances acquises par l'élève sur ces changements de représentation, mais nous nous interrogeons aussi sur la comparaison de ces deux items : les deux items sont-ils indépendants (c'est-à-dire, est-ce que le pourcentage de réussite à l'item 6 est le même à l'item 3 pour tous les participants au test, qu'ils aient ou non réussi ou échoué l'item 3) ?

Pour répondre à cette question, nous avons choisi de regrouper nos données dans un tableau croisé des réussites et des échecs à chacun des deux items (Annexe 4, tableau 2). Les élèves qui réussissent à transformer une écriture fractionnaire en une écriture décimale (item 3) et qui échouent à la transformation d'écriture inverse (item 6) sont plus nombreux que ceux qui réalisent correctement la transformation d'un nombre d'une écriture décimale à une écriture fractionnaire en (item 6) et qui échouent à l'inverse (item 3). Plusieurs éléments peuvent expliquer ce résultat : l'un d'eux se rapporte à l'ordre habituel⁹ d'enseignement des nombres décimaux à partir des fractions, l'autre au fait qu'il y a une pluralité d'écritures fractionnaires correspondant à une écriture décimale ($3,2$ peut s'écrire $32/10$ ou $320/100$ ou $3+2/10$), alors que, pour une écriture fractionnaire donnée, une seule écriture décimale lui correspond.

⁹ Dans la majorité des manuels de cycle 3 consultés, la progression d'enseignement adoptée pour les fractions et les nombres décimaux est dans l'ordre pour le CM 1 : Fractions (fraction partage dans le cas général) – Fractions décimales – Nombres décimaux – Opérations avec les nombres décimaux.

3.4.3 Écritures décimales : degrés de certitude et qualité de la réponse

Adjoindre un degré de certitude à la réponse donnée permet d'apporter des éléments d'évaluation plus précis et de distinguer des états de connaissance spécifiés, tels que « l'ignorance reconnue (réponse incorrecte et peu sûre), la connaissance incomplète (réponse correcte mais peu sûre), la connaissance assurée (réponse correcte et très sûre), l'ignorance ignorée (réponse incorrecte et très sûre) » (Gilles, 1996). Pour les trois items (3, 4 et 6) évaluant des connaissances sur les nombres décimaux, nous avons mis en relation, dans le tableau 8, la qualité de la réponse avec le degré de certitude qui lui a été accordé.

	Item 3	Item 4	Item 6
Réponse juste et $0 \leq DC \leq 2$: connaissance ignorée	15,48%	9,68%	10,32%
Réponse juste et $3 \leq DC \leq 4$: connaissance assurée	36,77%	58,71%	32,26%
Réponse fausse et $0 \leq DC \leq 2$: ignorance reconnue	17,42%	12,26%	22,58%
Réponse fausse et $3 \leq DC \leq 4$: ignorance ignorée	20,00%	12,90%	26,45%

Tableau 8 – Répartition des élèves selon les items et selon les degrés de certitude accordés à leur réponse

Ce tableau permet de révéler que les élèves semblent plus incertains concernant leur aptitude à répondre à l'item 6 (passage d'une écriture à virgule à une écriture sous forme de fraction décimale) qu'à l'item 3 (passage d'une écriture sous forme de fraction décimale à une écriture à virgule) puisqu'ils sont à la fois moins assurés quand ils répondent correctement et à la fois plus réalistes concernant leur ignorance. Globalement, nous constatons qu'un élève sur cinq ignore son ignorance ; ce constat rejoint ce que nous avons observé sur l'« ignorance ignorée » des élèves « faibles » et sur le fait qu'environ 20% des élèves conçoivent l'écriture à virgule a,b comme celle de la fraction a/b.

Conclusion

Indubitablement, il existe des différences de stratégies de réponse aux QCM selon que l'on soit un adulte, étudiant au lycée ou à l'université, et un enfant, élève de l'école primaire. Ces différences sont, en partie, la conséquence d'une approche différenciée de l'évaluation dans les différentes institutions scolaires. Les adultes chercheraient davantage « la bonne réponse », quitte à convoquer le hasard en

dernier recours, alors que les enfants chercheraient davantage à trouver « une réponse ».

Nous avons constaté que les élèves de l'école primaire n'adoptaient pas un type de stratégies uniforme pour résoudre les tâches auxquelles ils étaient confrontés, mais qu'ils en changeaient suivant la nature de ces tâches. Les stratégies qu'ils utilisent dépendent également de leur niveau de connaissances en mathématiques. Ainsi, les élèves les plus faibles sont plus enclins à utiliser des stratégies de repli ou de substitution, même s'ils rechignent à choisir leur réponse totalement au hasard. Il semblerait qu'un contrat didactique s'opère dans le cadre de ce type d'évaluation, ne permettant pas aux élèves d'envisager cette stratégie comme scolairement acceptable, à moins qu'ils n'aient pas suffisamment de recul pour utiliser le hasard comme recours possible ; la notion de « contrat d'évaluation » développée par Van den Heuvel Panhuizen (2005) en perspective de celle de « contrat didactique » (Brousseau, 1998) mériterait, au vu des résultats de cette recherche, d'être approfondie dans le cas spécifique des questions sous forme de QCM à l'école primaire lorsqu'elles sont effectivement présentées en situation d'évaluation.

Les items relatifs au lien entre écriture décimale et fraction ont produit des résultats assez inquiétants car, au-delà des scores de réussite assez faibles témoignant d'une mauvaise compréhension des décimaux, ils nous ont alerté sur le fait que certains élèves utilisent des règles erronées qu'ils appliquent avec un degré de certitude élevé (ignorance ignorée). Ainsi, l'intégration du degré de certitude au choix d'une réponse nous a permis, à l'échelle de notre test et de la taille de l'échantillon, de mieux mesurer le degré de connaissance des élèves et l'ancrage de leurs conceptions erronées, surtout pour les fractions - décimaux.

Les travaux menés en didactologie sur l'utilisation des degrés de certitude et de réalisme avec des étudiants offrent des perspectives pour faire évoluer les évaluations diagnostiques et formatives pour des élèves de primaire. Même si l'utilisation de QCM en cours d'apprentissage n'est pas préconisée (Grégoire & Laveault, 1997), on mesure à travers ce travail et ce type de dispositif, l'intérêt que présentent de telles informations pour l'enseignant et pour l'élève. L'utilisation en classe, par l'enseignant de tels indicateurs, semble une voie intéressante pour mieux appréhender le niveau de ses élèves et pour réguler son enseignement. Pour les erreurs fréquentes repérées par les didacticiens, nous convenons avec Pluvinage, (1979) que « pointer cette erreur et faire remarquer son caractère attractif à un élève qui l'a commise est un service à lui rendre ». En ce sens, les degrés de réalisme peuvent être un levier intéressant pour faire entrer l'élève dans l'évaluation de ses propres apprentissages et considérer l'erreur comme faisant partie du processus d'apprentissage.

Enfin, même si cette recherche ne l'a pas pris en compte, nous avons pu constater l'aspect potentiellement formateur du QCM, notamment en situation de résolution

de problèmes, puisque la forme même de la question offre à l'élève de possibles rétroactions sur la qualité de son résultat et peut le conduire à reprendre ses calculs. Au-delà d'un apprentissage possible de stratégies propres aux QCM (écarter les réponses invraisemblables, utiliser les ordres de grandeur, tester les différentes propositions, etc.), apprendre avec les QCM pourrait permettre d'aider à la résolution de problèmes et développer chez l'élève une certaine autonomie : d'une part les différents choix de réponse « sécurisent » l'élève, mais d'autre part les choix de réponses lui permettent d'exercer un auto-contrôle sur ses résultats.

Ces questions nous semblent importantes à traiter car les représentations des enseignants sur leurs élèves, qu'ils soient performants ou en difficulté, ont une incidence non négligeable sur leurs pratiques. En effet, si un enseignant estime que certains élèves échouent par manque de confiance en eux, ils peuvent être amenés à leur proposer des situations mathématiquement moins riches qui produisent plus facilement des réussites (logique de réussite immédiate, Butlen, Peltier, Pézard, 2002), mais qui au final réduisent les apprentissages effectifs des élèves. Les travaux du groupe RESEIDA ont montré que les difficultés des élèves pouvaient être, en partie, due à une différenciation pédagogique inadaptée conçue par des enseignants ayant pourtant une réelle volonté d'aider leurs élèves (Rochex & Crinon, 2011). Les questions d'évaluations sont au cœur des pratiques enseignantes, et peuvent donc aider à mieux appréhender la réalité des connaissances des élèves. Nous espérons que cette étude autour des stratégies de réponses aux QCM et des degrés de certitude y contribuera.

Bibliographie

ADDA J. (1976), Difficultés liées à la présentation des questions en mathématiques, *Educational Studies in Mathematics* 7, 3–22.

BANDURA A. (2003), *Auto-efficacité. Le sentiment d'efficacité personnelle*, De Boeck, Bruxelles.

BLOOM B. S., HASTING J. T. & MADAUS G. F. (1971), *Handbook on formative and summative evaluation of student learning*, Mac Graw Hill, New York.

BOULE S. & LAVEAULT D. (2011), Utilisation du degré de certitude et de réalisme dans un contexte d'évaluation diagnostique, dans *Des mécanismes pour assurer la validité de l'interprétation de la mesure en éducation, Volume 2*, (Eds. Raïche & alii), 31-48, Presses de l'Université du Québec, Québec.

BOURNY G., CENS I, COPPENS N., DUPE C., FUMEL S., JOST R., MONNIER A-L., OBERT M-C., OLIVIER Y., PEYLET D., ROCHER T., TALAMONI C. (2003), *L'évaluation internationale PISA 2003 : compétences des élèves français en mathématiques, compréhension de l'écrit et sciences*, Les dossiers de

l'enseignement scolaire, n°180, Direction de l'évaluation de la prospective et de la performance, Paris.

BROUSSEAU G. (1998), *Théorie des situations didactiques*, La Pensée Sauvage, Grenoble.

BUTLEN D., PELTIER M-L., PEZARD M. (2002), Nommé(s) en REP, comment font-ils ? Pratiques de professeurs des écoles enseignant les mathématiques en ZEP : cohérence et contradictions, *Revue française de pédagogie* **140**, 41-52.

COMITI C. & NEYRET R. (1979), A propos des problèmes rencontrés lors de l'enseignement des décimaux en classe de CM, *Grand N* **18**, 5-20.

CHOPPIN B.H. (1975), Guessing the answer on objective tests, *British Journal of Educational Psychology* **45**, 206-213.

DE LANDSHEERE G. (1979), *Dictionnaire de l'évaluation et de la recherche en éducation*, Presses universitaires de France, Paris.

DOUADY R. & PERRIN-GLORIAN M.J (1986), *Liaison école – collège : Nombres décimaux*. Brochure n°62, IREM de Paris 7, Paris.

DUVAL R., PLUVINAGE F. (1977), Démarches individuelles de réponse en mathématique, *Educational Studies in Mathematics* **8-1**, 51-116.

GILLES J-L. (1996), Utilisation des degrés de certitude et normes de réalisme en situation d'examen et d'auto-estimation à FA.P.S.E. - ULG, *Colloque de l'ADMEE-EUROPE "Dix années de travaux de recherche en évaluation"*.

GREGOIRE J., LAVEAULT D. (1997), *Introduction aux théories des tests en sciences humaines*, De Boeck, Bruxelles.

HUNT, D. (1993). Human self assessment - Theory and application to learning and testing, dans *NATO ASI Series, Item Banking : Interactive Testing and Self Assessment*, (Eds D. Leclercq, J. Bruno), **112**, 177-189, Springer Verlag, Berlin.

KATZ, I., BENNETT, R. E., & BERGER, A. (2000), Effects of Response Format on Difficulty of SAT Mathematics Items: It's Not the Strategy. *Journal of Educational Measurement* **37-1**, 39-57.

LECLERCQ D. (1986), *La conception des questions à choix multiple*, Labor, Bruxelles.

LECLERCQ D. (1987), *Qualité des questions et signification des scores*, Labor, Bruxelles.

LECLERCQ D. &POUMAY M. (2004), Une définition opérationnelle de la métacognition et ses mises en œuvre. *Communication présentée à la 21ème conférence internationale de l'AIPU*, Marrakech.

LECLERCQ D. (2006), L'évolution des QCM, dans *Recherche sur l'évaluation en éducation* (Eds. Figari & Mottier-Lopez), 139-146, L'Harmattan, Paris.

PLUVINAGE F. (1979), Loto-questionnaires (pour l'évaluation et l'auto-contrôle en mathématiques), *Educational Studies in Mathematics* **10**, 443-485.

ROCHEX J. & CRINON J. (2011), *La construction des inégalités scolaires. Au cœur des pratiques et des dispositifs d'enseignement*, Presses Universitaires de Rennes, Rennes.

RODITI E. (2007), La comparaison des nombres décimaux, conception et expérimentation d'une aide aux élèves en difficulté, *Annales de Didactique et de Sciences Cognitives* **12**, 55-81.

RONDIER M. (2004), « A. Bandura. Auto-efficacité. Le sentiment d'efficacité personnelle », *L'orientation scolaire et professionnelle* **33/3**, 475-476.

SIMKIN, M. G., & KUECHLER, W. L. (2005). Multiple Choice Tests and Student Understanding: What is the Connection? *Decision Sciences Journal of Innovative Education* **3-1**, 73-97.

VAN DEN HEUVEN-PANHUIZEN, M. (2005). The role of contexts in assessment problems in mathematics, *For the Learning of Mathematics* **25(2)**, 2-9.

NATHALIE SAYAC & NADINE GRAPIN

Université Paris-Est, Laboratoire de Didactique André Revuz

(LDAR-EA 4434)

nathalie.sayac@u-pec.fr

nadine.grapin@u-pec.fr

Annexe 1 : Énoncé des sept QCM pour les élèves**Item 1 :**

Dans une heure, combien y a-t-il de secondes ?

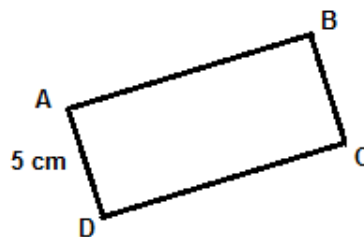
- 360 secondes
 3 060 secondes
 3 600 secondes
 6 300 secondes

1	2	3	4
pas sûr du tout	pas très sûr	sûr	sûr et certain

Item 2

ABCD est un rectangle dont la largeur est 5 cm.

Quelle doit être la longueur de ce rectangle pour que son périmètre soit égal à 30 cm ?



- 6 cm
 10 cm
 25 cm
 14 cm

1	2	3	4
pas sûr du tout	pas très sûr	sûr	sûr et certain

Item 3

A quel nombre correspond la fraction $\frac{62}{10}$?

- 6,2
 0,62
 62,10
 620

1	2	3	4
pas sûr du tout	pas très sûr	sûr	sûr et certain

Item 4

Lorsqu'on divise 872 par 100, on obtient:

- 0,872
 8,72
 87,20
 87 200

1	2	3	4
pas sûr du tout	pas très sûr	sûr	sûr et certain

Item 5

Sophia va à la boulangerie et achète 3 croissants à 1,20 € l'un.

Elle paie avec un billet de 10 €. Combien lui rendra caissière ?

- 7,20 €
 3,60 €
 8,80 €
 6,40 €

1	2	3	4
pas sûr du tout	pas très sûr	sûr	sûr et certain

Item 6

Quelle est la fraction égale à 237,8 ?

$\frac{237}{8}$

$\frac{2378}{100}$

$\frac{2378}{10}$

$\frac{237}{10}$

1	2	3	4
pas sûr du tout	pas très sûr	sûr	sûr et certain

Item 7

Pour faire une bonne confiture, il faut mettre 300 g de sucre pour 400 g de fruits.
 Combien faut-il mettre de sucre pour 1 kg de fruits ?

- 600 g
 700 g
 750 g
 800 g

1	2	3	4
pas sûr du tout	pas très sûr	sûr	sûr et certain

Annexe 2 : Explicitation des distracteurs pour chaque item

Item 1 (Item extrait de l'évaluation Cedre 2008)

Dans une heure, combien y a-t-il de secondes ?

- 360 secondes 3 060 secondes 3 600 secondes 6 300 secondes

Les distracteurs choisis correspondent à des réponses erronées proches, en termes d'écriture, de la bonne réponse. 360 peut correspondre aussi à une erreur de calcul du produit de 60 x 60 si un élève ne connaît pas de façon automatique la réponse.

Item 2

ABCD est un rectangle dont la largeur est 5 cm. Quelle doit être la longueur de ce rectangle pour que son périmètre soit égal à 30 cm ?

- 6 cm : *correspond à la confusion de la formule du périmètre avec celle de l'aire.*
- 10 cm : *bonne réponse.*
- 25 cm : *30 cm - 5 cm*
- 14 cm : *double confusion : utilisation de la formule de l'aire d'un rectangle pour déterminer la longueur (30 : 5 = 6) puis utilisation du résultat avec la formule du périmètre d'un rectangle pour calculer (à nouveau) la longueur : 30 cm - 2 x 5 cm - 6 cm = 14 cm.*

Notons que la question était accompagnée d'une figure à l'échelle (cf annexe 1) ; certains élèves ont trouvé 10 cm en constatant sur la figure que la longueur mesurait le double de la largeur.

Item 3

A quel nombre correspond la fraction $\frac{62}{10}$?

- 6,2 : *bonne réponse.*
- 0,62 : *confusion dixième – centième ou technique de division par 100 non maîtrisée.*
- 62,10 : *le trait de fraction est perçu comme un séparateur, au même titre que la virgule.*
- 620 : *confusion dixièmes / dizaines ou confusion diviser par 10 / multiplier par 10.*

Item 4 (Item extrait de l'évaluation Cedre 2008)

Lorsqu'on divise 872 par 100, on obtient:

- 0,872 8,72 87,20 87 200

Les mauvaises réponses correspondent à des résultats provenant d'un usage non maîtrisé d'une technique de la division par 10 : décalage de la virgule de 3 rangs pour 0,872 ; décalage de 1 rang pour 87,20 et pour 87200 confusion entre une division par 100 et une multiplication par 100.

Item 5

Sophia va à la boulangerie et achète 3 croissants à 1,20 € l'un.

Elle paie avec un billet de 10 €. Combien lui rendra caissière ?

- 7,20 € : *double erreur de calcul : $3 \times 1,20\text{€} = 3,20\text{€}$ et $10\text{€} - 3,20\text{€} = 7,20\text{€}$.*
 3,60 € : *correspond au prix des croissants (résultat d'une étape intermédiaire)*
 8,80 € : *$10\text{€} - 1,20\text{€}$ (l'élève a pris en compte le prix d'un seul croissant)*
 6,40 € : *bonne réponse*

Item 6 (Item extrait de l'évaluation Cedre 2008)

Quelle est la fraction égale à 237,8 ?

- $\frac{237}{8}$: *le trait de fraction est perçu comme un séparateur, au même titre que la virgule.*
 $\frac{2378}{100}$: *technique non maîtrisée de la division par 100 (on garde 3 chiffres avant la virgule).*
 $\frac{2378}{10}$: *bonne réponse*
 $\frac{237}{10}$: *la virgule, perçue comme un séparateur et la présence d'un chiffre après la virgule conduit l'élève à penser à une division par 10.*

Item 7

Pour faire une bonne confiture, il faut mettre 300 g de sucre pour 400 g de fruits. Combien faut-il mettre de sucre pour 1 kg de fruits ?

- 600 g : *$2 \times 300\text{g}$ ou $1\ 000\text{g} - 400\text{g}$*
 700 g : *$300\text{g} + 400\text{g}$*
 750 g : *bonne réponse*
 800 g : *écart de 100g entre 300g et 400g qui s'applique sur 1000g ou $2 \times 400\text{g}$.*

Annexe 3

Tableaux récapitulatifs des stratégies utilisées suivant le groupe d'élèves (faibles, moyens, forts) et selon l'item.

<i>Faibles</i>	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7
Stratégies A	64,5%	83,9%	87,1%	61,3%	54,8%	77,4%	25,8%
Stratégies B	16,1%	16,1%	12,9%	35,5%	19,4%	16,1%	61,3%
Stratégies C	19,4%	0,0%	0,0%	3,2%	25,8%	6,5%	12,9%

<i>Moyens</i>	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7
Stratégies A	73,0%	84,1%	54,0%	77,8%	65,1%	55,6%	34,9%
Stratégies B	17,5%	7,9%	36,5%	20,6%	4,8%	33,3%	52,4%
Stratégies C	9,5%	7,9%	9,5%	1,6%	30,2%	11,1%	12,7%

<i>Forts</i>	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7
Stratégies A	81,0%	98,3%	74,1%	93,1%	69,0%	58,6%	58,6%
Stratégies B	12,1%	1,7%	22,4%	3,4%	5,2%	19,0%	25,9%
Stratégies C	6,9%	0,0%	3,4%	3,4%	25,9%	22,4%	15,5%

Annexe 4

Tableau croisé des réponses aux items 3 et 6 :

Item 6 \ Item 3	1 : 237/8	2 : 2378/100	3 : 2378/10	4 : 237/10
1 : 6,2	13,5%	5,2%	34,8%	1,3%
2 : 0,62	9,7%	0,6%	5,8%	0,0%
3 : 62,10	18,1%	0,6%	1,9%	1,3%
4 : 620	1,3%	1,9%	3,9%	0,0%

Tableau croisé des réussites et des échecs aux items 3 et 6 (en nombre d'élèves et en pourcentage) :

Item 6 \ Item 3	Echecs	Réussites	Total
Réussites	31 (20%)	54 (35 %)	85 (55 %)
Echecs	52 (33 %)	18 (12 %)	70 (45 %)
Total	83 (53 %)	72 (47 %)	155