

**SYMPORIUM - L'ÉVALUATION DE TÂCHES COMPLEXES EN MATHÉMATIQUES :
QUELS OUTILS POUR UN MEILLEUR DIAGNOSTIC ?**

Géry Marcoux* & Annick Fagnant **

(Organisateurs du symposium)

Isabelle Demonty^{3*} & Virginie Dupont^{4*}, Christophe Dierendonck^{5*} & Annick Fagnant,
Nathalie Loyer^{6*}, Géry Marcoux* et Marc Vantourout^{7*}**

(Communicants)

Marcel Crahay^{8*} & Bernard Rey^{9*}

(discutants)

* Université de Genève (Suisse) - Gery.Marcoux@unige.ch

** Université de Liège(Belgique) – afagnant@ulg.ac.be

^{3*} Université de Liège (Belgique) - isabelle.demonty@ulg.ac.be

^{4*} Université de Liège (Belgique) - virginie.dupont@ulg.ac.be

^{5*} Université du Luxembourg (Luxembourg) – christophe.dierendonck@uni.lu

^{6*}Université de Montréal (Canada) - nathalie.loye@umontreal.ca

^{7*} Université Paris Descartes (France) - marc.vantourout@parisdescartes.fr

^{8*} Universités de Genève (Suisse) et de Liège (Belgique) - Marcel.Crahay@unige.ch

^{9*} Université Libre de Bruxelles (Belgique) - brey@ulb.ac.be

Mots-clés : Tâches complexes – Mathématiques – Evaluation diagnostique

Résumé. La notion de compétence occupe une place importante dans le champ scolaire de la majorité des pays francophones. En mathématiques, il est reconnu que la résolution de problèmes constitue en soi l'expression singulière d'une compétence. Le symposium comportait cinq communications, issues d'autant de pays francophones (Belgique, Canada, France, Luxembourg, Suisse) qui, par des focus différents, ont tenté d'affiner les modes de compréhension et d'explication des difficultés éprouvées par les élèves face à la résolution de tâches complexes en mathématiques. Le présent texte reprend le cadrage introductif du symposium et un résumé de chacune des interventions.

1. Cadrage du symposium (Géry Marcoux & Annick Fagnant)

La notion de compétence occupe une place importante dans le champ scolaire de la majorité des pays francophones. De nombreux courants existants en Sciences de l'Education s'en sont emparés avec pour conséquence actuelle une diversité foisonnante de définitions et de conceptions (Coulet, 2011). Dès lors, si la diversité des recherches a sans conteste un intérêt certain, les moments de synthèse sur un sujet ponctuel sont tout aussi essentiels.

En mathématiques, mais aussi plus largement dans le monde de la recherche en Sciences de l'Education, il est reconnu que la résolution de problèmes constitue en soi l'expression singulière d'une compétence (Crahy, Verschaffel, De Corte & Grégoire, 2005). Ce sera le premier point d'ancre de nos échanges.

Dans ce domaine plus spécifique, les références épistémologiques restent toutefois nombreuses et renvoient à des approches différentes avec des méthodologies diverses sur des objets riches et

multiples. Aussi, est-il nécessaire de cadrer d'avantage le propos en déterminant avec plus de précisions l'enjeu de ce symposium.

Sur la base de références diverses (didactique des mathématiques, psychologie cognitive, psychodidactique des maths, psychologie ergonomique, psychopédagogie,...) propres aux ancrages scientifiques justifiés par chacun des communicants et de leurs choix méthodologiques (analyse de tâches, analyse de productions, questionnaires d'enquête, hiérarchisation d'attributs, utilisation du modèle de Rasch, analyses par cluster,...), nous aborderons la question essentielle : quels outils pour un meilleur diagnostic en mathématiques ?

Ces deux termes (outil et diagnostic) ne sont évidemment pas dus au hasard mais servent de fil rouge à la réflexion. Dans le cas présent, le terme "outil" est à prendre au sens de Gilles (1978) comme "une matière préparée pour l'usage qu'on veut en faire, une forme raisonnée" (p.143). Ce construit, la réflexion sur son élaboration, sur ses implications et ses conséquences sont par essence à destination des enseignants. De même, le terme "diagnostic" est à comprendre au sens partiel que lui donnent Rey, Carette, Defrance et Khan (2003), c'est-à-dire "un processus qui vise à identifier les relations éventuelles entre facteurs (quelle qu'en soit la nature) susceptibles d'avoir une incidence privilégiée sur l'apprentissage (individuel ou collectif)" (p.39). Dans ce cas précis, il semble que les chercheurs ont un rôle important à jouer afin d'identifier ces facteurs, leurs relations et incidences et ainsi permettre, à terme, aux acteurs de l'enseignement d'être capable par eux-mêmes de les discerner (renvoyant de fait au sens étymologique grec du terme "diagnostic").

Le texte qui suit présente le résumé des cinq communications dans l'ordre dans lequel elles ont été présentées lors du symposium. Chacune d'elle, par des focus différents, a tenté d'affiner les modes de compréhension et d'explication des difficultés éprouvées par les élèves face à la résolution de tâches complexes en mathématiques.

2. Un diagnostic oui, mais lequel ? (Nathalie Loyer, Université de Montréal)

La présente étude s'intéresse à la perception des enseignants relativement aux besoins et aux difficultés rencontrées par les élèves en mathématique dans les centres de formation professionnelle de la grande région de Montréal au Québec. En effet, ces derniers sont actuellement aux prises avec un taux d'échec et d'abandon élevé, particulièrement dans les programmes incluant des compétences et des contenus mathématiques. Ainsi, cette étude constitue le point de départ à l'élaboration d'épreuves en mathématique ayant le pouvoir de diagnostiquer les besoins et les difficultés afin, dans un deuxième temps, de proposer des solutions pour aider les élèves risquant de décrocher.

Lorsqu'un sujet répond à un item dans un test ou résout un problème ou une tâche complexe en mathématique, il doit posséder une ou plusieurs habiletés, des connaissances, des savoir-faire et éventuellement des stratégies, que nous pourrions étiqueter sous le dénominateur générique commun d'attributs. Dans une approche diagnostique, ce sont alors les niveaux de maîtrise ou d'appropriation de ces attributs qui sont dignes d'intérêt, plutôt que la note finale du sujet. Il est pourtant souvent difficile d'observer ces attributs directement, surtout lorsque le nombre de sujets qui passent l'épreuve est grand, et encore plus lorsque l'instrument utilisé est constitué, en tout ou en partie, d'items à choix de réponses pour faire face à la correction d'un grand nombre de copies. Intuitivement, il semble possible de déduire la présence ou l'absence de ces attributs à partir des réponses ou des actions des sujets. À l'échelle de la salle de classe, cette déduction peut être le fruit d'une approche analytique basée sur l'observation des traces laissées sur la copie, ou de l'élève en action. Elle est alors réalisée au cas par cas, notamment en utilisant des critères d'évaluation dans des grilles descriptives détaillées. Sur une plus grande échelle, lorsque l'évaluation est réalisée avec un test, des modèles statistiques de classification diagnostique MCD ont été développés (voir par ex. Loyer, 2010; Roussos, et al., 2007; Tatsuoka, 2009) afin d'estimer le degré de maîtrise d'attributs reliés aux items d'un test, à partir du patron de réponses de chaque sujet.

Pour élaborer des tests ayant le pouvoir de réaliser un diagnostic relativement à une liste d'attributs, la démarche doit respecter plusieurs étapes (Loye, et al., 2011) dont la première consiste à définir quels attributs devraient faire l'objet d'un diagnostic. S'il existe une littérature abondante sur le sujet dans le domaine de la mathématique (Par exemple, Birenbaum, et al., 1993; Birenbaum, et al., 2004; Hartz, 2002; Loyer, 2008; Milewski & Baron, 2002), la participation des enseignants des centres de formation professionnelle est essentielle pour cibler, dans les divers programmes, les attributs qu'ils jugent utiles de diagnostiquer précocement chez leurs élèves afin d'être en mesure de leur offrir une remédiation ou une formation adaptée efficaces, et de les amener à résoudre correctement des problèmes complexes.

Les participants à cette étude sont des enseignants en mathématique, ou dont la discipline repose sur des notions mathématiques, ainsi que des conseillers pédagogiques travaillant dans plusieurs centres de formation professionnelle de la grande région de Montréal. Les données sont récoltées à l'aide d'un questionnaire en ligne. La liste des attributs incluse dans ce questionnaire provient d'une analyse systématique des attributs issus de la littérature (Loye, 2008). Le questionnaire a de plus été révisé par deux enseignantes en mathématique qui participent régulièrement à l'élaboration des tests de sélection en mathématique pour plusieurs centres de formation professionnelle de Montréal.

Le questionnaire comporte trois ensembles de questions. Le premier ensemble vise à récolter des informations contextuelles sur la situation professionnelle du répondant (fonction, cours, programme, type de clientèle); le second s'intéresse à la perception générale des répondants relativement aux besoins et aux difficultés des élèves en mathématique. Enfin, le troisième porte sur le niveau perçu des difficultés des élèves en lien avec les divers contenus mathématiques (par exemple : concepts et opérations de base sur des fractions) et sur l'utilité de diagnostiquer les faiblesses des élèves relativement à une liste d'attributs (par exemple : utiliser des figures, tableaux, graphiques ; comprendre les problèmes posés sous forme de phrases, etc.). Une question ouverte sur la description des besoins en matière de test diagnostique en mathématique clôture le questionnaire.

Les données récoltées font l'objet d'une analyse descriptive, puis sont analysées à l'aide d'un modèle de Rash. L'objectif est de placer sur une même échelle de mesure les divers contenus et attributs mathématiques. Les réponses fournies à la question ouverte sont traitées séparément. Les résultats prennent la forme d'une hiérarchisation des attributs et des contenus mathématiques et constituent ainsi le point de départ à une démarche d'élaboration d'épreuves diagnostiques en mathématique adaptées aux besoins du milieu de la formation professionnelle.

3. Situations complexes en mathématiques : que proposent les diverses approches figurant dans la littérature pour évaluer des compétences ? (Marc Vantourout, Université Paris Descartes)

1. Pour répondre à l'objectif proposé aux participants du symposium, il nous semble que l'analyse des situations, tâches ou épreuves (peu importe la terminologie) soumises aux élèves constitue une entrée intéressante.

En adoptant le point de vue de la psychologie ergonomique, nous appréhendons l'enseignant-évaluateur comme un pourvoyeur de tâches complexes (quand il s'agit de compétences) d'enseignement et d'évaluation. Fabriquer, transformer ou utiliser tout simplement des situations constitue l'une des deux principales facettes de l'activité de l'évaluateur, l'autre étant leur évaluation ou correction (là aussi, peu importe la terminologie).

Nous formulons les questions abordées lors de la communication ainsi : Peut-on utiliser les situations complexes que l'on trouve dans la littérature scientifique pour construire des évaluations en mathématiques ? En quoi et à quelles conditions leur utilisation permet-elle d'accéder à une meilleure compréhension des difficultés des élèves ? Notons que ce questionnement véhicule l'hypothèse implicite selon laquelle ces situations pourraient aider les enseignants à réaliser leur activité d'évaluateur.

Le rattachement à la thématique du colloque repose sur une idée largement partagée dans la littérature : l'indissociabilité des notions de « compétences » et de « tâches/situations complexes ».

2.1. La première partie de la communication sera plutôt théorique sans toutefois négliger la dimension empirique. Nous présenterons les approches psycho-didactiques en mathématiques (APDM, voir 2.2) dont relèvent nos travaux dans le domaine de l'évaluation. Nous les confronterons à d'autres approches dont nous rappellerons les principales finalités afin de ne pas tomber dans des débats inutiles. Notons que certaines de ces approches, qui s'ignorent actuellement, auraient beaucoup à tirer les unes des autres car elles traitent toutes finalement des processus d'apprentissage et d'enseignement de contenus disciplinaires identifiés (nous pensons, par exemple, à la didactique des mathématiques (DDM) qui semble « ignorer » les travaux des psychopédagogues belges). Nous aimeraisons pouvoir également contribuer à des rapprochements bénéfiques.

La confrontation s'organisera autour d'aspects et préoccupations partagés ou non par les diverses approches : l'analyse des tâches, l'analyse *a priori* des situations, leur complexité, sens, authenticité, réalisme, dimension concrète ainsi que la question de la mathématisation et modélisation. C'est en « psycho-didacticien » convaincu de l'utilité de certains concepts et résultats empiriques de la DDM dans le domaine de l'évaluation – donc avec des critères inspirés de nos préoccupations – que nous mettrons en parallèle les approches retenues. Outre les approches (APDM, DDM) déjà mentionnés, nous nous intéresserons aux propositions de Rey, Carette et coll. (2003), des psychopédagogues belges (Crahay, Verschaffel & coll., 2005) ainsi qu'à celles des méthodologues canadiens et belges (par exemple, Scallion, 2004).

2.2. Dans ce paragraphe, nous indiquons quelques éléments théoriques et réflexions qui seront présentés lors de la communication. Avec Maury (2002) qui se fonde sur des arguments de nature épistémologique et pragmatique, nous faisons figurer dans les APDM des auteurs tels Vergnaud (1983, 1990), Levain (1997), Brun (1979, 1994) et Julo (1995). Ceux-ci, en réalisant des travaux fondés empiriquement, ont montré que « l'apport des psychologies cognitives est indiscutable, au moins en ce qu'elles nous éclairent sur les rapports personnels aux savoirs ou conception des sujets » (Maury, 2002). Les APDM s'inscrivent dans une perspective fonctionnaliste et postulent que l'étude du fonctionnement cognitif du sujet individuel en situation est essentielle pour l'enseignement et l'évaluation (Vantourout & Maury, 2006). La notion de conception est ici centrale car elle intègre une analyse du savoir (objectif) et des connaissances (parfois implicites) du sujet en situation. Les APDM, grâce aux analyses de la tâche et *a priori* (empruntées respectivement à la psychologie et à la DDM) complétées par des analyses des productions et procédures d'élèves, permettent de construire des situations qui ont vraiment du sens du point de vue des mathématiques (Fabre, 1999) et d'appréhender des questions liées à la modélisation et à la régulation. Notons que d'autres (Rey & coll., 2003) adoptent partiellement, au moins au niveau « théorique », des positions extrêmement proches des nôtres et de celles des psychopédagogues belges.

3. Dans la deuxième partie, qui sera la plus importante lors de la communication, nous nous centrerons sur quelques situations complexes tirées de travaux renvoyant aux approches mentionnées. Deux questions essentielles orientent les analyses en cours : celle de la validité des contenus et celle d'une véritable évaluation des compétences (voir le paragraphe précédent). L'analyse intègre également des éléments auxquels « tiennent » les auteurs des différentes approches (par exemple, l'authenticité, le réalisme) au risque d'en « négliger » d'autres (les connaissances en jeu, l'anticipation des procédures, la validité des contenus). Pour constituer un corpus centré sur des contenus identifiés (APDM obligent), nous avons retenu des situations, pouvant être proposées à des élèves de 5^e-7^e années (CM2-5^e en France), où est impliquée la proportionnalité. Certaines de ces situations ont été soumises à des sujets afin de disposer d'éléments relatifs à leur résolution et de confirmations quant à la validité de leurs contenus. Un premier résultat, très général, concerne la nécessité de bien distinguer les situations à orientation « psychologique » et « pédagogiques », qui ne prennent pas toujours en compte les contenus, et celles à orientation « didactique ».

4. Pour la partie discussion, nous nous appuierons sur des points qui orientent nos travaux passés et actuels.

Nous pensons que les APDM fournissent un cadre adéquat pour une véritable évaluation des compétences. En effet, elles permettent, au-delà d'un simple intérêt pour la performance, la réalisation d'une analyse extrêmement fine des compétences dans leur dimension cognitive (Vergnaud, 2001 ; Crahay, 2006). Elles pourraient favoriser le développement de pratiques de régulations reposant sur une meilleure compréhension des productions d'élèves. Les APDM pourraient également contribuer à doter les enseignants d'outils de régulation afin qu'ils s'engagent dans une évaluation à visée formative de situations complexes en mathématiques (Brun, 1979 ; Vantourout, 2004 et 2007).

4. La mesure des compétences des élèves face à la résolution d'une tâche complexe en algèbre élémentaire : l'apport d'indicateurs issus des recherches centrées sur les difficultés spécifiques des élèves en algèbre (Isabelle Demonty & Virginie Dupont, Université de Liège)

Depuis quelques années déjà, la Communauté française de Belgique propose aux enseignants de mathématiques des outils d'évaluation des compétences élaborés en référence au modèle d'évaluation des compétences développé par Rey et al. (2003). Ce modèle se décline en trois phases d'évaluation représentant chacune un degré de compétences. La première phase permet d'évaluer le troisième degré de compétences, assimilé aux « compétences complexes » : l'élève doit « savoir choisir et combiner correctement plusieurs compétences élémentaires pour traiter une situation nouvelle et complexe » (Carette, 2007a). La deuxième phase d'évaluation propose aux élèves de résoudre la même tâche complexe, découpée cette fois en tâches élémentaires. Ici, ce sont les compétences du deuxième degré qui sont évaluées. Enfin, la dernière phase correspond à l'évaluation d'une série de tâches simples décontextualisées. On évalue ainsi le premier degré de compétences qui correspond aux « compétences élémentaires » aussi appelées « procédures ».

Cette approche permet aux enseignants de voir dans quelle mesure des lacunes apparaissent à l'un des trois niveaux suivants: se situent-elles au niveau des procédures (compétences de 1^{er} degré), de la mise en œuvre de celles-ci dans des situations cadrées (compétences de 2^e degré) ou de l'exploration complète et autonome d'une situation inédite (compétences de 3^e degré) ? L'approche proposée rend cependant peu compte des difficultés spécifiques des élèves face à la résolution d'un problème mobilisant des compétences algébriques élémentaires. De nombreuses études se sont intéressées à cette problématique (voir Demonty, 2005 pour une synthèse des travaux menés dans ce domaine). La modélisation de problèmes algébriques est souvent assimilée à une traduction d'un énoncé en langage mathématique. Pour Julo (1995), « cette manière d'analyser la démarche de modélisation est celle du mathématicien qui sait résoudre le problème et ne correspond pas à ce qui se passe au niveau des processus cognitifs » (p. 65). Schmidt et Bednarz (2002) précisent cette idée dans le contexte des problèmes algébriques. Selon ces auteurs, lorsqu'un élève appréhende un tel problème selon une logique arithmétique, il mobilise des grandeurs connues présentées dans l'énoncé et organise sur cette base une séquence d'opérations qui aboutira, en fin de parcours à la réponse du problème. Il procède ainsi du connu vers l'inconnu, en s'appuyant constamment sur des repères empruntés au contexte. Dans l'approche algébrique, la démarche de pensée est très différente : le raisonnement s'appuie dès le départ sur une quantité inconnue souvent représentée par un substitut symbolique. A partir de celui-ci, l'élève reconstruit, sous la forme statique de l'équation, les relations stipulées dans le problème. Par la suite, l'équation est résolue, sans qu'il n'y ait à ce moment de repères empruntés en contexte.

Si, dans la tâche complexe, l'élève est libre d'utiliser l'approche de son choix, la décomposition du problème en sous-problème le plonge d'emblée dans une approche algébrique du problème, semblable selon les termes de Julo (1995) « à celle du mathématicien qui sait déjà résoudre le problème ». Comment l'élève parvient-il à tirer profit de cet éventuel changement d'approche de la situation ? Quels indicateurs fournir aux enseignants pour les aider à mieux comprendre les difficultés spécifiques des élèves face à ce type de problèmes ? Telles sont les questions abordées

dans notre communication. Elle se propose de présenter et de discuter des résultats découlant de passation d'un outil centré sur la résolution d'un problème d'algèbre élémentaire. Cet outil a été soumis auprès de 167 élèves de 3^e année de l'enseignement secondaire général issus de 8 écoles en Communauté française de Belgique. Ces élèves ont répondu à un questionnaire en deux phases distinctes : 20 minutes pour résoudre la tâche complexe « le cadeau collectif » issue des outils d'évaluation et 50 minutes pour résoudre les 2^e (tâche décomposée) et 3^e (procédures) parties de l'outil. Le choix de tester cet outil d'évaluation destiné aux élèves du premier degré de l'enseignement secondaire auprès d'élèves de 3^e secondaire s'explique par le fait que nous voulions garantir que cette matière ait été travaillée dans toutes les classes. Par la suite, les épreuves ont été corrigées en respectant les critères fournis par les concepteurs de l'épreuve.

Si l'épreuve centrée sur les procédures élémentaire est réussie par une majorité d'élèves, il n'en va pas de même pour les deux autres épreuves. Certains élèves réussissent même mieux la tâche complexe que sa décomposition en sous-problèmes. Comment expliquer ce paradoxe ? Un retour à l'analyse des productions de ces élèves fait apparaître que bon nombre d'entre eux ont réussi à résoudre la tâche complexe selon une démarche arithmétique. Ils se sont avérés par la suite incapables de tirer profit de la décomposition du problème pour appréhender algébriquement cette fois la situation. Après avoir présenté les résultats découlant de cette analyse, notre intervention envisagera des indicateurs permettant aux enseignants d'appréhender les difficultés des élèves dans la transition entre une démarche arithmétique et algébrique de résolution de problèmes. Se pose également la question de l'intérêt de la décomposition du problème en sous-problèmes dans des contenus où manifestement diverses démarches de réflexion peuvent être mise en œuvre pour résoudre la tâche complexe sans que des procédures algébriques évaluées dans la troisième partie de l'outil n'aient dû être mobilisées.

5. Evaluer des compétences en mathématiques dans le cadre d'une épreuve externe à large échelle : tâches complexes, tâches décomposées et tâches simples : quel pouvoir informatif ? (Christophe Dierendonck, Université du Luxembourg & Annick Fagnant, Université de Liège)

La notion de « compétence » a pris une place considérable dans la plupart des référentiels européens. On attend aujourd'hui des élèves qu'ils soient capables de mobiliser, de façon intégrée, différentes ressources pour accomplir une tâche, généralement qualifiée de complexe (Carette, 2007 ; De Ketela & Gérard, 2005). Si l'intérêt de mettre l'accent sur la « mobilisation » est généralement partagé, force est de constater que les élèves éprouvent d'importantes difficultés face à ces tâches complexes.

Le modèle « en phases » proposé par Rey et al. (2003) et le « modèle » d'évaluation des ressources impliquées dans les tâches complexes au départ de tâches élémentaires indépendantes proposé par Crahay et Detheux (2005) sont deux approches visant à mieux comprendre les difficultés éprouvées par les élèves face aux tâches complexes. Le premier objectif de l'étude que nous avons menée est de voir dans quelle mesure il est possible de combiner ces deux démarches d'évaluation et d'analyser leur pouvoir informatif complémentaire.

Un autre débat actuel en matière d'évaluation concerne les « divergences » entre les nouveaux modèles d'évaluation des compétences et les évaluations externes, nationales ou internationales, « classiques » qui, selon plusieurs auteurs, n'évaluerait pas réellement des compétences (Carette, 2007a,b ; De Ketela & Gérard, 2005).

Si l'on part d'une définition des compétences acceptant un niveau de « complexité » des tâches comme nécessitant « l'identification, la mobilisation et l'intégration de plus d'une procédure (...), une interprétation (ou un cadrage) de la situation et une organisation de la démarche de résolution » (Dierendonck & Fagnant, 2010, p. 13, définition actualisée), il est alors possible d'insérer des évaluations de compétences au sein d'une épreuve externe « classique » dans la

mesure où elles ne seront pas trop « consommatrices » en temps et pourront coexister avec d'autres types d'items d'évaluation. C'est à ce niveau que se situe le deuxième objectif de l'étude.

Dans le cadre des évaluations externes à large échelle menées en octobre 2009 et 2010 en troisième secondaire au Luxembourg, nous avons expérimenté un dispositif d'évaluation distinguant quatre types de tâches visant à apporter des informations diagnostiques : des tâches complexes de référence, des tâches complexes décomposées, des tâches élémentaires décontextualisées et des tâches élémentaires présentées dans un autre complexe problématique. En plus de ces tâches-cibles, les épreuves comprenaient une variété de tâches évaluant des compétences de 1^{er}, de 2^e ou de 3^e degré de façon à couvrir le plus largement possible (en fonction des contraintes temporelles, deux heures de cours de 50 minutes au maximum) les domaines de contenu investigués (à savoir les « nombres et opérations » et les « figures du plan et de l'espace »).

Dans les deux épreuves, la présence d'un nombre suffisant d'items (et leur indépendance) a autorisé l'utilisation d'un modèle de réponse à l'item (le modèle de Rash) permettant d'analyser, sur une échelle commune, le degré de difficulté des tâches et le niveau de compétence des élèves. Les résultats, corroborés dans les deux études, permettent de tirer trois constats principaux : (a) on n'observe pas de hiérarchie stricte entre les quatre types de tâches, (b) les tâches décomposées présentent un niveau de difficulté assez proche de celui des tâches complexes et (c) de nombreuses tâches élémentaires en contexte présentent un niveau de difficulté moindre que celui des tâches décomposées. Si le guidage et la décomposition des tâches complexes posait déjà question en matière d'enseignement (on n'apprend pas aux élèves à développer un « savoir-mobiliser » en faisant l'analyse de la situation à leur place et en leur indiquant le type de ressources adéquates), nos résultats conduisent à questionner son apport réel sur le plan diagnostique. Par ailleurs, dans la mesure où l'information fournie par l'évaluation des compétences de 2^e degré dans des situations indépendantes les unes des autres ne conduit pas aux mêmes constats que l'évaluation de ces « mêmes » compétences de 2^e degré dans des tâches décomposées, une épreuve « en phases » qui n'évaluerait ces compétences de 2^e degré que dans des tâches décomposées conduirait à sous-estimer, pour un nombre non négligeable d'élèves, le niveau de maîtrise de ces compétences.

Par ailleurs, une analyse complémentaire croisant les résultats obtenus par les élèves aux quatre types de tâches proposées démontre assez nettement que les élèves se comportent différemment de ce que l'on attend a priori d'un modèle diagnostique en phases supposées hiérarchiques. La diversité des résultats observés plaide pour l'intérêt d'épreuves constituées d'un nombre suffisant d'items et d'items suffisamment variés pour avoir le plus de chances de « saisir » les compétences réelles des élèves. Bien que des analyses complémentaires soient nécessaires pour mieux cerner le réel pouvoir diagnostique des épreuves proposées, les premiers résultats nous semblent témoigner de l'intérêt du modèle utilisé (combinant les apports de Rey et al. 2003 et de Crahay & Detheux, 2005) et, *a contrario*, du « danger » à évaluer les compétences uniquement dans quelques tâches complexes, même décomposées selon un modèle « en phases ».

Le texte complet de cette communication est accessible dans les Actes du 24^e colloque de l'Admée-Europe.

6. Modèle dérivé du modèle en 3 phases de Rey, Carette, Khan et Defrance : un pouvoir explicatif augmenté ? (Géry Marcoux, Université de Genève)

A l'heure actuelle, dans le monde scolaire, évaluer les compétences des élèves reste problématique. La difficulté réside peut-être, en partie, comme l'ont déjà souligné Rey, Carette, Defrance et Khan (2003), dans le fait que l'expression même d'une compétence nécessite de manière confondue l'utilisation d'automatismes et la capacité d'adaptation à la nouveauté. Dans la pratique enseignante, si l'enseignement à l'acquisition d'une procédure, en terme d'action standardisée, et son évaluation peuvent sembler ne pas présenter de difficultés insurmontables, il

en va tout autrement de la compétence ! En effet, dans l'évaluation (comme dans l'enseignement) d'une compétence, l'élément central n'est plus la capacité à répondre de façon adéquate à une forme de signal prédéterminé (une consigne explicite qui attend une action standardisée précise), mais la capacité du sujet d'analyser une situation "inédite" afin de choisir dans son répertoire des procédures qu'il pense adéquates (quitte à devoir les adapter) et d'organiser leur exécution selon un ordre a priori non défini. Autrement dit, se contenter d'indicateurs dichotomiques du type "fait" ou "non fait" suivant un modèle linéaire unique de résolution n'aurait qu'un intérêt relatif pour l'enseignant et/ou l'évaluateur. Ce qui importe plus, c'est l'observation et l'analyse du/des processus de résolution utilisé(s) et son/leur adéquation à la tâche proposée dans une visée scolaire.

Comment rendre compte de la non utilisation ou de l'utilisation entière ou partielle ou encore erronée de ces processus par le sujet ? Une première possibilité consiste à confronter le sujet à une tâche complexe et à analyser ses traces écrites pour observer voire inférer les processus utilisés et ainsi aller au-delà du simple constat. Une autre consiste à proposer plusieurs prises d'informations sur un même objet ; ce que proposent les modèles à phases tels que celui de Rey *et al.* (2003). Dans ce cas, si on adopte une position radicale, le constat ne se base plus sur les inférences issues de l'analyse des traces à disposition, mais sur des mises en relation entre des observables dans les différentes phases. Ainsi, pour le modèle en 3 phases de Rey *et al.* : la phase 1 (où le sujet est confronté à une tâche complexe qui exige la mobilisation d'un nombre significatif de procédures et d'éléments de savoir qu'il est censé posséder) se borne à établir un degré de réalisation ou d'avancement du sujet dans la réalisation de la tâche ; la phase 2 (où le sujet est confronté à la même tâche complexe mais avec étayage) procède de l'observation de l'utilisation ou non de procédures à un niveau inférieur (dans des tâches partielles qui composent la tâche complexe) ; la phase 3 (où le sujet est confronté à une batterie de procédures de bases impliquées dans la tâche globale et censées automatisées) constate la maîtrise ou non de ces procédures de base par une cotation binaire sur base d'une série d'items (réussi ou non). Il s'agit alors de comparer ce que le sujet a fait dans les 3 phases. L'inférence ne se fait plus alors sur la base de l'analyse des traces dans les différentes phases mais sur le rapport entre ces phases.

Un intérêt du modèle en phases réside donc sur le plus grand nombre d'observables sur lesquels l'enseignant et/ou l'évaluateur peut se baser. Toutefois, à notre sens, ce modèle n'empêche pas les inférences, et plus encore, il propose, nous semble-t-il, des indicateurs qui ne permettent parfois qu'une explication faible et peu robuste des processus évoqués. Tenant compte des apports récents concernant la résolution de problèmes (e.g. : Barouillet, P. & Camos, V. 2006 ; Clément, 2009 ; Crahay, Verschaeffel, De Corte & Grégoire, 2005), nous avons dès lors repris l'outil en 3 phases de Rey *et al.* pour en proposer un dérivé tant du point de vue de la conception des phases que de la correction de celles-ci. Sur la base d'une étude menée dans le courant 2009-2010 auprès d'élèves de l'enseignement obligatoire du canton de Genève [âge moyen : 10 ans], nous présentons ce modèle dérivé et nous discutons ses avantages et limites, la pertinence de nouveaux indicateurs (distinction faits déclaratifs-habiletés, indices de compréhension type1-type2,...) ainsi que leur portée concernant la compréhension fine des processus de résolution de problèmes arithmétiques.

Références

- Barouillet, P. & Camos, V. (2006). *La cognition mathématique chez l'enfant*. Marseille : Solal.
- Birenbaum, M., Kelly, A. E., & Tatsuoka, K. K. (1993). Diagnosing knowledge states in algebra using the rule-space model. *Journal for Research in Mathematics Education*, 24(5), 442-459.
- Birenbaum, M., Tatsuoka, C., & Yamada, T. (2004). Diagnostic assessment in TIMSS-R : Between-countries and within country comparisons of eighth graders' mathematics performance. *Studies in Educational Evaluation*, 30, 151-173.
- Brun, J. (1979). L'évaluation formative dans un enseignement différencié de mathématiques. In L. Allal, J. Cardinet & P. Perrenoud (Eds.), *L'évaluation formative dans un enseignement différencié*, (pp. 203-215). Berne : Peter Lang.

*Actes du 24^e colloque de l'Admée-Europe
L'évaluation des compétences en milieu scolaire et en milieu professionnel*

- Brun, J. (1994). Evolution des rapports entre la psychologie du développement cognitif et la didactique des mathématiques. In M. Artigue, R. Gras, C. Laborde & P. Tavignot (Eds.), *Vingt ans de didactique des mathématiques en France – Hommage à Guy Brousseau et Gérard Vergnaud*, (pp. 67-83). Grenoble : La Pensée Sauvage éditions.
- Carette, V. (2007a). Les implications de la notion de compétence sur l'évaluation. *Education-Formation*, 286, 51-61.
- Carette, V. (2007b). L'évaluation au service de la gestion des paradoxes liés à la notion de compétence. *Mesure et évaluation*. 30(2), 49-71.
- Crahay, M. (2006). Dangers, incertitudes et incomplétude de la logique de la compétence en éducation. *Revue française de pédagogie*, 154, 97-110.
- Crahay, M. & Dethieux, M. (2005). L'évaluation des compétences, une entreprise impossible ? (Résolution de problèmes complexes et maîtrise de procédures mathématiques). *Mesure et évaluation en éducation*, 28 (1), p. 57-76.
- Crahay, M., Verschaffel, L., De Corte, E. & Grégoire, J. (Eds.).(2005). *Enseignement et apprentissage des mathématiques. Que disent les recherches psychopédagogiques*. Bruxelles : De Boeck & Larcier.
- Clément E. (2009). *La Résolution de problèmes : à la découverte de la flexibilité cognitive*. Paris : A. Colin.
- Coulet, J.-C. (2011). La notion de compétence : un modèle pour décrire, évaluer et développer les compétences. *Le Travail humain*, 74(1), 1-30.
- Demonty, I. (2005). La transition entre l'arithmétique et l'algèbre dans le contexte de la résolution de problèmes. In M. Crahay, L. Verschaffel, E. De Corte & J. Grégoire (Eds.), *Enseignement et apprentissage des mathématiques. Que disent les recherches psychopédagogiques ?* (151-170). Bruxelles : De Boeck.
- De Ketela, J.-M. & Gérard, F.-M. (2005). La validation des épreuves d'évaluation selon l'approche par les compétences, *Mesure et évaluation en éducation*, 28(3), 1-26.
- Dierendonck, C. & Fagnant, A. (2010). Quelques réflexions autour des épreuves d'évaluation développées dans le cadre de l'approche par compétences. *Le Bulletin de l'ADMEE-EUROPE*, 2010/1, 5-20.
- Everitt, B. (1993). *Cluster analysis*. New York : Wiley.
- Fabre, M. (1999). Situations-problèmes et savoir scolaire. Paris : Presses Universitaires de France.
- Gille, B. (1978). *Histoire des techniques*. Paris : Gallimard.
- Hartz, S. M. (2002). *A Bayesian framework for the unified model for assessing cognitive abilities: Blending theory with practicality*. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign, Urbana-Champaign, IL.
- Julo, J. (1995). Représentation des problèmes et réussite en mathématiques. Un apport de la psychologie cognitive à l'enseignement. Rennes : presses universitaires.
- Levain, J.-P. (1997). Faire des maths autrement – Développement cognitif et proportionnalité. Paris : L'Harmattan.
- Loye, N. (2008). *Conditions d'élaboration de la matrice Q des modèles cognitifs et impact sur sa validité et sa fidélité*. Unpublished doctoral thesis, University of Ottawa, Ottawa.
- Loye, N. (2010). 2010, odyssée des modèles de classification diagnostique (MCD). *Mesure et évaluation en éducation*, 33(3), 75-98.
- Loye, N., Caron, F., Pineault, J., Tessier-Baillargeon, M., Burney-Vincent, C., & Gagnon, M. (2011). La validité du diagnostic issu d'un mariage entre didactique et mesure sur un test existant. Dans G. Raîche, K. Paquette-Côté et D. Magis (Eds), *Des mécanismes pour assurer la validité de l'interprétation de la mesure en éducation, volume 2* (pp. 11-30). Ste-Foy, Québec: Presses de l'Université du Québec.
- Maury, S. (2001). Didactique des mathématiques et psychologie cognitive : un regard comparatif sur trois approches psychologiques. *Revue française de pédagogie*, 137, 85-93.

*Actes du 24^e colloque de l'Admée-Europe
L'évaluation des compétences en milieu scolaire et en milieu professionnel*

- Milewski, G. B., & Baron, P. A. (2002). *Extending DIF methods to inform aggregate reports on cognitive skills*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, New Orleans.
- Rey, B., Carette, V., Defrance, A. & Kahn, S. (2003). *Les compétences à l'école : apprentissage et évaluation*, Bruxelles : De Boeck.
- Roussos, L., DiBello, L., Stout, W., Hartz, S., Henson, R. A., & Templin, J. (2007). The fusion model skills diagnosis system. In J. P. Leighton & M. J. Gierl (Eds.), *Cognitive diagnostic assessment for education: Theory and applications* (pp. 275-318). New York, NY: Cambridge University Press.
- Scallon, G. (2004). L'évaluation des apprentissages dans une approche par compétences. Bruxelles : De Boeck.
- Schmidt, S & Bednarz, N. (2002). Arithmetical and algebraic types of reasoning used by pre-service teachers in a problem-solving context. *Canadian Journal for Mathematics, Service and Technology Education*, 2(1), p.67-90.
- Tatsuoka, K. K. (2009). *Cognitive assessment: An introduction to the rule space method*. New York: Routledge Taylor & Francis Group.
- Vantourout, M. (2004). Etude de l'activité et des compétences de professeurs des écoles et de professeurs de mathématiques dans des situations « simulées » d'évaluation à visée formative en mathématiques. Thèse de doctorat non publiée. Université Paris V.
- Vantourout, M. & Maury, S. (2006). Quelques résultats relatifs aux connaissances disciplinaires de professeurs stagiaires dans des situations simulées d'évaluation de productions d'élèves en mathématiques. *Revue des sciences de l'éducation*, vol. 32/3, 759-782.
- Vantourout, M. (2007). Étude de l'activité évaluative de professeurs stagiaires confrontés à des productions d'élèves en mathématiques : quel référent pour l'évaluateur ? *Mesure et évaluation en éducation*, 30(3), 29-58.
- Vergnaud, G. (1983). Introduction. *Recherches en didactique des mathématiques*, Vol. 4/1, 9-25.
- Vergnaud, G. (1990). La théorie des champs conceptuels. *Recherches en didactique des mathématiques*, Vol. 10/2.3, 133-170.
- Vergnaud, G. (2001). Psychologie du développement cognitif et évaluation des compétences. In Figari, G & Achouche, M. (Eds.), *L'activité évaluative réinterrogée*, (pp. 43-51). Bruxelles : De Boeck.