

**SYMPOSIUM - 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 Loye^{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'Éducation 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'Éducation, il est reconnu que la résolution de problèmes constitue en soi l'expression singulière d'une compétence (Crahay, Verschaffel, De Corte & Grégoire, 2005). Ce sera le premier point d'ancrage 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 n  cessaire de cadrer d'avantage le propos en d  terminant avec plus de pr  cisions l'enjeu de ce symposium.

Sur la base de r  f  rences diverses (didactique des math  matiques, psychologie cognitive, psychodidactique des maths, psychologie ergonomique, psychop  dagogie,...) propres aux ancrages scientifiques justifi  s par chacun des communicants et de leurs choix m  thodologiques (analyse de t  ches, analyse de productions, questionnaires d'enqu  te, hi  rarchisation d'attributs, utilisation du mod  le de Rasch, analyses par cluster,...), nous aborderons la question essentielle : quels outils pour un meilleur diagnostic en math  matiques ?

Ces deux termes (outil et diagnostic) ne sont   videmment pas dus au hasard mais servent de fil rouge    la r  flexion. Dans le cas pr  sent, le terme "outil" est    prendre au sens de Gille (1978) comme "une mati  re pr  par  e pour l'usage qu'on veut en faire, une forme raisonn  e" (p.143). Ce construit, la r  flexion sur son   laboration, sur ses implications et ses cons  quences sont par essence    destination des enseignants. De m  me, le terme "diagnostic" est    comprendre au sens partiel que lui donnent Rey, Carette, Defrance et Khan (2003), c'est-   dire "un processus qui vise    identifier les relations   ventuelles entre facteurs (quelle qu'en soit la nature) susceptibles d'avoir une incidence privil  gi  e sur l'apprentissage (individuel ou collectif)" (p.39). Dans ce cas pr  cis, il semble que les chercheurs ont un r  le important    jouer afin d'identifier ces facteurs, leurs relations et incidences et ainsi permettre,    terme, aux acteurs de l'enseignement d'  tre capable par eux-m  mes de les discerner (renvoyant de fait au sens   tymologique grec du terme "diagnostic").

Le texte qui suit pr  sente le r  sum   des cinq communications dans l'ordre dans lequel elles ont   t   pr  sent  es lors du symposium. Chacune d'elles, par des focus diff  rents, a tent   d'affiner les modes de compr  hension et d'explication des difficult  s   prouv  es par les   l  ves face    la r  solution de t  ches complexes en math  matiques.

2. Un diagnostic oui, mais lequel ? (Nathalie Loye, Universit   de Montr  al)

La pr  sente   tude s'int  resse    la perception des enseignants relativement aux besoins et aux difficult  s rencontr  es par les   l  ves en math  matique dans les centres de formation professionnelle de la grande r  gion de Montr  al au Qu  bec. En effet, ces derniers sont actuellement aux prises avec un taux d'  chec et d'abandon   lev  , particuli  rement dans les programmes incluant des comp  tences et des contenus math  matiques. Ainsi, cette   tude constitue le point de d  part    l'  laboration d'  preuves en math  matique ayant le pouvoir de diagnostiquer les besoins et les difficult  s afin, dans un deuxi  me temps, de proposer des solutions pour aider les   l  ves risquant de d  crocher.

Lorsqu'un sujet r  pond    un item dans un test ou r  sout un probl  me ou une t  che complexe en math  matique, il doit poss  der une ou plusieurs habilet  s, des connaissances, des savoir-faire et   ventuellement des strat  gies, que nous pourrions   tiqueter sous le d  nominateur g  n  rique commun d'attributs. Dans une approche diagnostique, ce sont alors les niveaux de ma  trise ou d'appropriation de ces attributs qui sont dignes d'int  r  t, plut  t 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'  preuve est grand, et encore plus lorsque l'instrument utilis   est constitu  , en tout ou en partie, d'items    choix de r  ponses pour faire face    la correction d'un grand nombre de copies. Intuitivement, il semble possible de d  duire la pr  sence ou l'absence de ces attributs    partir des r  ponses ou des actions des sujets.    l'  chelle de la salle de classe, cette d  duction peut   tre le fruit d'une approche analytique bas  e sur l'observation des traces laiss  es sur la copie, ou de l'  l  ve en action. Elle est alors r  alis  e au cas par cas, notamment en utilisant des crit  res d'  valuation dans des grilles descriptives d  taill  es. Sur une plus grande   chelle, lorsque l'  valuation est r  alis  e avec un test, des mod  les statistiques de classification diagnostique MCD ont   t   d  velopp  s (voir par ex. Loye, 2010; Roussos, et al., 2007; Tatsuoka, 2009) afin d'estimer le degr   de ma  trise d'attributs reli  s aux items d'un test,    partir du patron de r  ponses de chaque sujet.

Pour   laborer des tests ayant le pouvoir de r  aliser un diagnostic relativement    une liste d'attributs, la d  marche doit respecter plusieurs   tapes (Loye, et al., 2011) dont la premi  re consiste    d  finir quels attributs devraient faire l'objet d'un diagnostic. S'il existe une litt  rature abondante sur le sujet dans le domaine de la math  matique (Par exemple, Birenbaum, et al., 1993; Birenbaum, et al., 2004; Hartz, 2002; Loye, 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 pr  cocement chez leurs   l  ves afin d'  tre en mesure de leur offrir une r  m  diation ou une formation adapt  e efficaces, et de les amener    r  soudre correctement des probl  mes complexes.

Les participants    cette   tude sont des enseignants en math  matique, ou dont la discipline repose sur des notions math  matiques, ainsi que des conseillers p  dagogiques travaillant dans plusieurs centres de formation professionnelle de la grande r  gion de Montr  al. Les donn  es sont r  colt  es    l'aide d'un questionnaire en ligne. La liste des attributs incluse dans ce questionnaire provient d'une analyse syst  matique des attributs issus de la litt  rature (Loye, 2008). Le questionnaire a de plus   t   r  vis   par deux enseignantes en math  matique qui participent r  guli  rement    l'  laboration des tests de s  lection en math  matique pour plusieurs centres de formation professionnelle de Montr  al.

Le questionnaire comporte trois ensembles de questions. Le premier ensemble vise    r  colter des informations contextuelles sur la situation professionnelle du r  pondant (fonction, cours, programme, type de client  le) ; le second s'int  resse    la perception g  n  rale des r  pondants relativement aux besoins et aux difficult  s des   l  ves en math  matique. Enfin, le troisi  me porte sur le niveau per  u des difficult  s des   l  ves en lien avec les divers contenus math  matiques (par exemple : concepts et op  rations de base sur des fractions) et sur l'utilit   de diagnostiquer les faiblesses des   l  ves relativement    une liste d'attributs (par exemple : utiliser des figures, tableaux, graphiques ; comprendre les probl  mes pos  s sous forme de phrases, etc.). Une question ouverte sur la description des besoins en mati  re de test diagnostique en math  matique cl  ture le questionnaire.

Les donn  es r  colt  es font l'objet d'une analyse descriptive, puis sont analys  es    l'aide d'un mod  le de Rash. L'objectif est de placer sur une m  me   chelle de mesure les divers contenus et attributs math  matiques. Les r  ponses fournies    la question ouverte sont trait  es s  par  ment. Les r  sultats prennent la forme d'une hi  rarchisation des attributs et des contenus math  matiques et constituent ainsi le point de d  part    une d  marche d'  laboration d'  preuves diagnostiques en math  matique adapt  es aux besoins du milieu de la formation professionnelle.

3. Situations complexes en math  matiques : que proposent les diverses approches figurant dans la litt  rature pour   valuer des comp  tences ? (Marc Vantourout, Universit   Paris Descartes)

1. Pour r  pondre    l'objectif propos   aux participants du symposium, il nous semble que l'analyse des situations, t  ches ou   preuves (peu importe la terminologie) soumises aux   l  ves constitue une entr  e int  ressante.

En adoptant le point de vue de la psychologie ergonomique, nous appr  hendons l'enseignant-  valuateur comme un pourvoyeur de t  ches complexes (quand il s'agit de comp  tences) d'enseignement et d'  valuation. Fabriquer, transformer ou utiliser tout simplement des situations constitue l'une des deux principales facettes de l'activit   de l'  valuateur, l'autre   tant leur   valuation ou correction (l   aussi, peu importe la terminologie).

Nous formulons les questions abord  es lors de la communication ainsi : Peut-on utiliser les situations complexes que l'on trouve dans la litt  rature scientifique pour construire des   valuations en math  matiques ? En quoi et    quelles conditions leur utilisation permet-elle d'acc  der    une meilleure compr  hension des difficult  s des   l  ves ? Notons que ce questionnement v  hicule l'hypoth  se implicite selon laquelle ces situations pourraient aider les enseignants    r  aliser leur activit   d'  valuateur.

Le rattachement    la th  matique du colloque repose sur une id  e largement partag  e dans la litt  rature : l'indissociabilit   des notions de « comp  tences » et de « t  ches/situations complexes ».

2.1. La premi  re partie de la communication sera plut  t th  orique sans toutefois n  gliger la dimension empirique. Nous pr  senterons les approches psycho-didactiques en math  matiques (APDM, voir 2.2) dont rel  vent nos travaux dans le domaine de l'  valuation. Nous les confronterons    d'autres approches dont nous rappellerons les principales finalit  s afin de ne pas tomber dans des d  bats inutiles. Notons que certaines de ces approches, qui s'ignorent actuellement, auraient beaucoup    tirer les unes des autres car elles traitent toutes finalement des processus d'apprentissage et d'enseignement de contenus disciplinaires identifi  s (nous pensons, par exemple,    la didactique des math  matiques (DDM) qui semble « ignorer » les travaux des psychop  dagogues belges). Nous aimerions pouvoir   galement contribuer    des rapprochements b  n  fiques.

La confrontation s'organisera autour d'aspects et pr  occupations partag  s ou non par les diverses approches : l'analyse des t  ches, l'analyse a priori des situations, leur complexit  , sens, authenticit  , r  alisme, dimension concr  te ainsi que la question de la math  matisation et mod  lisation. C'est en « psycho-didacticien » convaincu de l'utilit   de certains concepts et r  sultats empiriques de la DDM dans le domaine de l'  valuation – donc avec des crit  res inspir  s de nos pr  occupations – que nous mettrons en parall  le les approches retenues. Outre les approches (APDM, DDM) d  j   mentionn  es, nous nous int  resserons aux propositions de Rey, Carette et coll. (2003), des psychop  dagogues belges (Crahay, Verschaffel & coll., 2005) ainsi qu'   celles des m  thodologues canadiens et belges (par exemple, Scallon, 2004).

2.2. Dans ce paragraphe, nous indiquons quelques   l  ments th  oriques et r  flexions qui seront pr  sent  s lors de la communication. Avec Maury (2002) qui se fonde sur des arguments de nature   pist  mologique 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 r  alisant des travaux fond  s empiriquement, ont montr   que « l'apport des psychologies cognitives est indiscutable, au moins en ce qu'elles nous   clairent sur les rapports personnels aux savoirs ou conception des sujets » (Maury, 2002). Les APDM s'inscrivent dans une perspective fonctionnaliste et postulent que l'  tude du fonctionnement cognitif du sujet individuel en situation est essentielle pour l'enseignement et l'  valuation (Vantourout & Maury, 2006). La notion de conception est ici centrale car elle int  gre une analyse du savoir (objectif  ) et des connaissances (parfois implicites) du sujet en situation. Les APDM, gr  ce aux analyses de la t  che et a priori (emprunt  es respectivement    la psychologie et    la DDM) compl  t  es par des analyses des productions et proc  dures d'  l  ves, permettent de construire des situations qui ont vraiment du sens du point de vue des math  matiques (Fabre, 1999) et d'appr  hender des questions li  es    la mod  lisation et    la r  gulation. Notons que d'autres (Rey & coll., 2003) adoptent partiellement, au moins au niveau « th  orique », des positions extr  mement proches des n  tres et de celles des psychop  dagogues belges.

3. Dans la deuxi  me partie, qui sera la plus importante lors de la communication, nous nous centrerons sur quelques situations complexes tir  es de travaux renvoyant aux approches mentionn  es. Deux questions essentielles orientent les analyses en cours : celle de la validit   des contenus et celle d'une v  ritable   valuation des comp  tences (voir le paragraphe pr  c  dent). L'analyse int  gre   galement des   l  ments auxquels « tiennent » les auteurs des diff  rentes approches (par exemple, l'authenticit  , le r  alisme) au risque d'en « n  gliger » d'autres (les connaissances en jeu, l'anticipation des proc  dures, la validit   des contenus). Pour constituer un corpus centr   sur des contenus identifi  s (APDM obligent), nous avons retenu des situations, pouvant   tre propos  es    des   l  ves de 5^e-7^e ann  es (CM2-5^e en France), o   est impliqu  e la proportionnalit  . Certaines de ces situations ont   t   soumises    des sujets afin de disposer d'  l  ments relatifs    leur r  solution et de confirmations quant    la validit   de leurs contenus. Un premier r  sultat, tr  s g  n  ral, concerne la n  cessit   de bien distinguer les situations    orientation « psychologique » et « p  dagogiques », qui ne prennent pas toujours en compte les contenus, et celles    orientation « didactique ».

4. Pour la partie discussion, nous nous appuyons sur des points qui orientent nos travaux pass  s et actuels.

Nous pensons que les APDM fournissent un cadre ad  quat pour une v  ritable   valuation des comp  tences. En effet, elles permettent, au-del   d'un simple int  r  t pour la performance, la r  alisation d'une analyse extr  mement fine des comp  tences dans leur dimension cognitive (Vergnaud, 2001 ; Crahay, 2006). Elles pourraient favoriser le d  veloppement de pratiques de r  gulations reposant sur une meilleure compr  hension des productions d'  l  ves. Les APDM pourraient   galement contribuer    doter les enseignants d'outils de r  gulation afin qu'ils s'engagent dans une   valuation    vis  e formative de situations complexes en math  matiques (Brun, 1979 ; Vantourout, 2004 et 2007).

4. La mesure des comp  tences des   l  ves face    la r  solution d'une t  che complexe en alg  bre   l  mentaire : l'apport d'indicateurs issus des recherches centr  es sur les difficult  s sp  cifiques des   l  ves en alg  bre (Isabelle Demonty & Virginie Dupont, Universit   de Li  ge)

Depuis quelques ann  es d  j  , la Communaut   fran  aise de Belgique propose aux enseignants de math  matiques des outils d'  valuation des comp  tences   labor  s en r  f  rence au mod  le d'  valuation des comp  tences d  velopp   par Rey et al. (2003). Ce mod  le se d  cline en trois phases d'  valuation repr  sentant chacune un degr   de comp  tences. La premi  re phase permet d'  valuer le troisi  me degr   de comp  tences, assimil   aux « comp  tences complexes » : l'  l  ve doit « savoir choisir et combiner correctement plusieurs comp  tences   l  mentaires pour traiter une situation nouvelle et complexe » (Carette, 2007a). La deuxi  me phase d'  valuation propose aux   l  ves de r  soudre la m  me t  che complexe, d  coup  e cette fois en t  ches   l  mentaires. Ici, ce sont les comp  tences du deuxi  me degr   qui sont   valu  es. Enfin, la derni  re phase correspond    l'  valuation d'une s  rie de t  ches simples d  contextualis  es. On   value ainsi le premier degr   de comp  tences qui correspond aux « comp  tences   l  mentaires » aussi appel  es « proc  dures ».

Cette approche permet aux enseignants de voir dans quelle mesure des lacunes apparaissent    l'un des trois niveaux suivants: se situent-elles au niveau des proc  dures (comp  tences de 1^{er} degr  ), de la mise en   uvre de celles-ci dans des situations cadr  es (comp  tences de 2^e degr  ) ou de l'exploration compl  te et autonome d'une situation in  dite (comp  tences de 3^e degr  ) ? L'approche propos  e rend cependant peu compte des difficult  s sp  cifiques des   l  ves face    la r  solution d'un probl  me mobilisant des comp  tences alg  briques   l  mentaires. De nombreuses   tudes se sont int  ress  es    cette probl  matique (voir Demonty, 2005 pour une synth  se des travaux men  s dans ce domaine). La mod  lisation de probl  mes alg  briques est souvent assimil  e    une traduction d'un   nonc   en langage math  matique. Pour Julio (1995), « cette mani  re d'analyser la d  marche de mod  lisation est celle du math  maticien qui sait r  soudre le probl  me et ne correspond pas    ce qui se passe au niveau des processus cognitifs » (p. 65). Schmidt et Bednarz (2002) pr  cisent cette id  e dans le contexte des probl  mes alg  briques. Selon ces auteurs, lorsqu'un   l  ve appr  hende un tel probl  me selon une logique arithm  tique, il mobilise des grandeurs connues pr  sent  es dans l'  nonc   et organise sur cette base une s  quence d'op  rations qui aboutira, en fin de parcours    la r  ponse du probl  me. Il proc  de ainsi du connu vers l'inconnu, en s'appuyant constamment sur des rep  res emprunt  s au contexte. Dans l'approche alg  brique, la d  marche de pens  e est tr  s diff  rente : le raisonnement s'appuie d  s le d  part sur une quantit   inconnue souvent repr  sent  e par un substitut symbolique. A partir de celui-ci, l'  l  ve reconstruit, sous la forme statique de l'  quation, les relations stipul  es dans le probl  me. Par la suite, l'  quation est r  solue, sans qu'il n'y ait    ce moment de rep  res emprunt  s en contexte.

Si, dans la t  che complexe, l'  l  ve est libre d'utiliser l'approche de son choix, la d  composition du probl  me en sous-probl  me le plonge d'  mbl  e dans une approche alg  brique du probl  me, semblable selon les termes de Julio (1995) «    celle du math  maticien qui sait d  j   r  soudre le probl  me ». Comment l'  l  ve parvient-il    tirer profit de cet   ventuel changement d'approche de la situation ? Quels indicateurs fournir aux enseignants pour les aider    mieux comprendre les difficult  s sp  cifiques des   l  ves face    ce type de probl  mes ? Telles sont les questions abord  es

dans notre communication. Elle se propose de pr  senter et de discuter des r  sultats d  coulant de la passation d'un outil centr   sur la r  solution d'un probl  me d'alg  bre   l  mentaire. Cet outil a   t   soumis aupr  s de 167   l  ves de 3  me ann  e de l'enseignement secondaire g  n  ral issus de 8   coles en Communaut   fran  aise de Belgique. Ces   l  ves ont r  pondu    un questionnaire en deux phases distinctes : 20 minutes pour r  soudre la t  che complexe « le cadeau collectif » issue des outils d'  valuation et 50 minutes pour r  soudre les 2   (t  che d  compos  e) et 3   (proc  dures) parties de l'outil. Le choix de tester cet outil d'  valuation destin   aux   l  ves du premier degr   de l'enseignement secondaire aupr  s d'  l  ves de 3  me secondaire s'explique par le fait que nous voulions garantir que cette mati  re ait   t   travaill  e dans toutes les classes. Par la suite, les   preuves ont   t   corrig  es en respectant les crit  res fournis par les concepteurs de l'  preuve.

Si l'  preuve centr  e sur les proc  dures   l  mentaire est r  ussie par une majorit   d'  l  ves, il n'en va pas de m  me pour les deux autres   preuves. Certains   l  ves r  ussissent m  me mieux la t  che complexe que sa d  composition en sous-probl  mes. Comment expliquer ce paradoxe ? Un retour    l'analyse des productions de ces   l  ves fait appara  tre que bon nombre d'entre eux ont r  ussi    r  soudre la t  che complexe selon une d  marche arithm  tique. Ils se sont av  r  s par la suite incapables de tirer profit de la d  composition du probl  me pour appr  hender alg  briquement cette fois la situation. Apr  s avoir pr  sent   les r  sultats d  coulant de cette analyse, notre intervention envisagera des indicateurs permettant aux enseignants d'appr  hender les difficult  s des   l  ves dans la transition entre une d  marche arithm  tique et alg  brique de r  solution de probl  mes. Se pose   galement la question de l'int  r  t de la d  composition du probl  me en sous-probl  mes dans des contenus o   manifestement diverses d  marches de r  flexion peuvent   tre mise en   uvre pour r  soudre la t  che complexe sans que des proc  dures alg  briques   valu  es dans la troisi  me partie de l'outil n'aient d     tre mobilis  es.

5. Evaluer des comp  tences en math  matiques dans le cadre d'une   preuve externe    large   chelle : t  ches complexes, t  ches d  compos  es et t  ches simples : quel pouvoir informatif ? (Christophe Dierendonck, Universit   du Luxembourg & Annick Fagnant, Universit   de Li  ge)

La notion de « comp  tence » a pris une place consid  rable dans la plupart des r  f  rentiels europ  ens. On attend aujourd'hui des   l  ves qu'ils soient capables de mobiliser, de fa  on int  gr  e, diff  rentes ressources pour accomplir une t  che, g  n  ralement qualifi  e de complexe (Carette, 2007 ; De Ketele & G  rard, 2005). Si l'int  r  t de mettre l'accent sur la « mobilisation » est g  n  ralement partag  , force est de constater que les   l  ves   prouvent d'importantes difficult  s face    ces t  ches complexes.

Le mod  le « en phases » propos   par Rey et al. (2003) et le « mod  le » d'  valuation des ressources impliqu  es dans les t  ches complexes au d  part de t  ches   l  mentaires ind  pendantes propos   par Crahay et Detheux (2005) sont deux approches visant    mieux comprendre les difficult  s   prouv  es par les   l  ves face aux t  ches complexes. Le premier objectif de l'  tude que nous avons men  e est de voir dans quelle mesure il est possible de combiner ces deux d  marches d'  valuation et d'analyser leur pouvoir informatif compl  mentaire.

Un autre d  bat actuel en mati  re d'  valuation concerne les « divergences » entre les nouveaux mod  les d'  valuation des comp  tences et les   valuations externes, nationales ou internationales, « classiques » qui, selon plusieurs auteurs, n'  valueraient pas r  ellement des comp  tences (Carette, 2007a,b ; De Ketele & G  rard, 2005).

Si l'on part d'une d  finition des comp  tences acceptant un niveau de « complexit   » des t  ches comme n  cessitant « l'identification, la mobilisation et l'int  gration de plus d'une proc  dure (...), une interpr  tation (ou un cadrage) de la situation et une organisation de la d  marche de r  solution » (Dierendonck & Fagnant, 2010, p. 13, d  finition actualis  e), il est alors possible d'ins  rer des   valuations de comp  tences au sein d'une   preuve externe « classique » dans la

mesure o   elles ne seront pas trop « consommatrices » en temps et pourront coexister avec d'autres types d'items d'  valuation. C'est    ce niveau que se situe le deuxi  me objectif de l'  tude.

Dans le cadre des   valuations externes    large   chelle men  es en octobre 2009 et 2010 en troisi  me secondaire au Luxembourg, nous avons exp  rim  nt   un dispositif d'  valuation distinguant quatre types de t  ches visant    apporter des informations diagnostiques : des t  ches complexes de r  f  rence, des t  ches complexes d  compos  es, des t  ches   l  mentaires d  contextualis  es et des t  ches   l  mentaires pr  sent  es dans un autre complexe probl  matique. En plus de ces t  ches-cibles, les   preuves comprenaient une vari  t   de t  ches   valuant des comp  tences de 1^{er}, de 2^e ou de 3^e degr   de fa  on    couvrir le plus largement possible (en fonction des contraintes temporelles, deux heures de cours de 50 minutes au maximum) les domaines de contenu investigu  s (   savoir les « nombres et op  rations » et les « figures du plan et de l'espace »).

Dans les deux   preuves, la pr  sence d'un nombre suffisant d'items (et leur ind  pendance) a autoris   l'utilisation d'un mod  le de r  ponse    l'item (le mod  le de Rasch) permettant d'analyser, sur une   chelle commune, le degr   de difficult   des t  ches et le niveau de comp  tence des   l  ves. Les r  sultats, corrobor  s dans les deux   tudes, permettent de tirer trois constats principaux : (a) on n'observe pas de hi  rarchie stricte entre les quatre types de t  ches, (b) les t  ches d  compos  es pr  sentent un niveau de difficult   assez proche de celui des t  ches complexes et (c) de nombreuses t  ches   l  mentaires en contexte pr  sentent un niveau de difficult   moindre que celui des t  ches d  compos  es. Si le guidage et la d  composition des t  ches complexes posait d  j   question en mati  re d'enseignement (on n'apprend pas aux   l  ves    d  velopper un « savoir-mobiliser » en faisant l'analyse de la situation    leur place et en leur indiquant le type de ressources ad  quates), nos r  sultats conduisent    questionner son apport r  el sur le plan diagnostique. Par ailleurs, dans la mesure o   l'information fournie par l'  valuation des comp  tences de 2^e degr   dans des situations ind  pendantes les unes des autres ne conduit pas aux m  mes constats que l'  valuation de ces « m  mes » comp  tences de 2^e degr   dans des t  ches d  compos  es, une   preuve « en phases » qui n'  valuerait ces comp  tences de 2^e degr   que dans des t  ches d  compos  es conduirait    sous-estimer, pour un nombre non n  gligeable d'  l  ves, le niveau de ma  trise de ces comp  tences.

Par ailleurs, une analyse compl  mentaire croisant les r  sultats obtenus par les   l  ves aux quatre types de t  ches propos  es d  montre assez nettement que les   l  ves se comportent diff  remment de ce que l'on attend a priori d'un mod  le diagnostique en phases suppos  es hi  rarchiques. La diversit   des r  sultats observ  s plaide pour l'int  r  t d'  preuves constitu  es d'un nombre suffisants d'items et d'items suffisamment vari  s pour avoir le plus de chances de « saisir » les comp  tences r  elles des   l  ves. Bien que des analyses compl  mentaires soient n  cessaires pour mieux cerner le r  el pouvoir diagnostique des   preuves propos  es, les premiers r  sultats nous semblent t  moigner de l'int  r  t du mod  le utilis   (combinant les apports de Rey et al. 2003 et de Crahay & Detheux, 2005) et, *a contrario*, du « danger »      valuer les comp  tences uniquement dans quelques t  ches complexes, m  me d  compos  es selon un mod  le « en phases ».

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

6. Mod  le d  riv   du mod  le en 3 phases de Rey, Carette, Khan et Defrance : un pouvoir explicatif augment   ? (G  ry Marcoux, Universit   de Gen  ve)

A l'heure actuelle, dans le monde scolaire,   valuer les comp  tences des   l  ves reste probl  matique. La difficult   r  side peut-  tre, en partie, comme l'ont d  j   soulign   Rey, Carette, Defrance et Khan (2003), dans le fait que l'expression m  me d'une comp  tence n  cessite de mani  re confondue l'utilisation d'automatismes et la capacit   d'adaptation    la nouveaut  . Dans la pratique enseignante, si l'enseignement    l'acquisition d'une proc  dure, en terme d'action standardis  e, et son   valuation peuvent sembler ne pas pr  senter de difficult  s insurmontables, il

en va tout autrement de la comp  tence ! En effet, dans l'  valuation (comme dans l'enseignement) d'une comp  tence, l'  l  ment central n'est plus la capacit      r  pondre de fa  on ad  quate    une forme de signal pr  d  termin   (une consigne explicite qui attend une action standardis  e pr  cise), mais la capacit   du sujet d'analyser une situation "in  dite" afin de choisir dans son r  pertoire des proc  dures qu'il pense ad  quates (quitte    devoir les adapter) et d'organiser leur ex  cution selon un ordre a priori non d  fini. Autrement dit, se contenter d'indicateurs dichotomiques du type "fait" ou "non fait" suivant un mod  le lin  aire unique de r  solution n'aurait qu'un int  r  t relatif pour l'enseignant et/ou l'  valuateur. Ce qui importe plus, c'est l'observation et l'analyse du/des processus de r  solution utilis  (s) et son/leur ad  quation    la t  che propos  e dans une vis  e scolaire.

Comment rendre compte de la non utilisation ou de l'utilisation enti  re ou partielle ou encore erron  e de ces processus par le sujet ? Une premi  re possibilit   consiste    confronter le sujet    une t  che complexe et    analyser ses traces   crites pour observer voire inf  rer les processus utilis  s et ainsi aller au-del   du simple constat. Une autre consiste    proposer plusieurs prises d'informations sur un m  me objet ; ce que proposent les mod  les    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 inf  rences issues de l'analyse des traces    disposition, mais sur des mises en relation entre des observables dans les diff  rentes phases. Ainsi, pour le mod  le en 3 phases de Rey *et al.* : la phase 1 (o   le sujet est confront      une t  che complexe qui exige la mobilisation d'un nombre significatif de proc  dures et d'  l  ments de savoir qu'il est cens   poss  der) se borne      tablir un degr   de r  alisation ou d'avancement du sujet dans la r  alisation de la t  che ; la phase 2 (o   le sujet est confront      la m  me t  che complexe mais avec   tayage) proc  de de l'observation de l'utilisation ou non de proc  dures    un niveau inf  rieur (dans des t  ches partielles qui composent la t  che complexe) ; la phase 3 (o   le sujet est confront      une batterie de proc  dures de bases impliqu  es dans la t  che globale et cens  es automatis  es) constate la ma  trise ou non de ces proc  dures de base par une cotation binaire sur base d'une s  rie d'items (r  ussi ou non). Il s'agit alors de comparer ce que le sujet a fait dans les 3 phases. L'inf  rence ne se fait plus alors sur la base de l'analyse des traces dans les diff  rentes phases mais sur le rapport entre ces phases.

Un int  r  t du mod  le en phases r  side donc sur le plus grand nombre d'observables sur lesquels l'enseignant et/ou l'  valuateur peut se baser. Toutefois,    notre sens, ce mod  le n'emp  che pas les inf  rences, et plus encore, il propose, nous semble-t-il, des indicateurs qui ne permettent parfois qu'une explication faible et peu robuste des processus   voqu  s. Tenant compte des apports r  cents concernant la r  solution de probl  mes (e.g. : Barouillet, P. & Camos, V. 2006 ; Cl  ment, 2009 ; Crahay, Verschaeffel, De Corte & Gr  goire, 2005), nous avons d  s lors repris l'outil en 3 phases de Rey *et al.* pour en proposer un d  riv   tant du point de vue de la conception des phases que de la correction de celles-ci. Sur la base d'une   tude men  e dans le courant 2009-2010 aupr  s d'  l  ves de l'enseignement obligatoire du canton de Gen  ve [  ge moyen : 10 ans], nous pr  sentons ce mod  le d  riv   et nous discutons ses avantages et limites, la pertinence de nouveaux indicateurs (distinction faits d  claratifs-habilit  s, indices de compr  hension type1-type2,...) ainsi que leur port  e concernant la compr  hension fine des processus de r  solution de probl  mes arithm  tiques.

R  f  rences

- Barouillet, P. & Camos, V. (2006). *La cognition math  matique 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'  valuation formative dans un enseignement diff  renci   de math  matiques. In L. Allal, J. Cardinet & P. Perrenoud (Eds.), *L'  valuation formative dans un enseignement diff  renci  *, (pp. 203-215). Berne : Peter Lang.

- Brun, J. (1994). Evolution des rapports entre la psychologie du d  veloppement cognitif et la didactique des math  matiques. In M. Artigue, R. Gras, C. Laborde & P. Tavnnot (Eds.), *Vingt ans de didactique des math  matiques en France – Hommage    Guy Brousseau et G  rard Vergnaud*, (pp. 67-83). Grenoble : La Pens  e Sauvage   ditions.
- Carette, V. (2007a). Les implications de la notion de comp  tence sur l'  valuation. *Education-Formation*, 286, 51-61.
- Carette, V. (2007b). L'  valuation au service de la gestion des paradoxes li  s    la notion de comp  tence. *Mesure et   valuation*. 30(2), 49-71.
- Crahay, M. (2006). Dangers, incertitudes et incompl  tude de la logique de la comp  tence en   ducation. *Revue fran  aise de p  dagogie*, 154, 97-110.
- Crahay, M. & Detheux, M. (2005). L'  valuation des comp  tences, une entreprise impossible ? (R  solution de probl  mes complexes et ma  trise de proc  dures math  matiques). *Mesure et   valuation en   ducation*, 28 (1), p. 57-76.
- Crahay, M., Verschaffel, L., De Corte, E. & Gr  goire, J. (Eds.).(2005). *Enseignement et apprentissage des math  matiques. Que disent les recherches psychop  dagogiques*. Bruxelles : De Boeck & Larcier.
- Cl  ment E. (2009). *La R  solution de probl  mes :    la d  couverte de la flexibilit   cognitive*. Paris : A. Colin.
- Coulet, J.-C. (2011). La notion de comp  tence : un mod  le pour d  crire,   valuer et d  velopper les comp  tences. *Le Travail humain*, 74(1), 1-30.
- Demonty, I. (2005).La transition entre l'arithm  tique et l'alg  bre dans le contexte de la r  solution de probl  mes. In M. Crahay, L. Verschaffel, E. De Corte & J. Gr  goire (Eds.), *Enseignement et apprentissage des math  matiques. Que disent les recherches psychop  dagogiques ?* (151-170). Bruxelles : De Boeck.
- De Ketele, J.-M. & G  rard, F.-M. (2005). La validation des   preuves d'  valuation selon l'approche par les comp  tences, *Mesure et   valuation en   ducation*, 28(3), 1-26.
- Dierendonck, C. & Fagnant, A. (2010). Quelques r  flexions autour des   preuves d'  valuation d  velopp  es dans le cadre de l'approche par comp  tences. *Le Bulletin de l'ADMEE-EUROPE*, 2010/1, 5-20.
- Everitt, B. (1993). *Cluster analysis*. New York : Wiley.
- Fabre, M. (1999). Situations-probl  mes 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). Repr  sentation des probl  mes et r  ussite en math  matiques. Un apport de la psychologie cognitive    l'enseignement. Rennes : presses universitaires.
- Levain, J.-P. (1997). Faire des maths autrement – D  veloppement cognitif et proportionnalit  . Paris : L'Harmattan.
- Loye, N. (2008). *Conditions d'  laboration de la matrice Q des mod  les cognitifs et impact sur sa validit   et sa fid  lit  *. Unpublished doctoral thesis, University of Ottawa, Ottawa.
- Loye, N. (2010). 2010, odyss  e des mod  les de classification diagnostique (MCD). *Mesure et   valuation en   ducation*, 33(3), 75-98.
- Loye, N., Caron, F., Pineault, J., Tessier-Baillargeon, M., Burney-Vincent, C., & Gagnon, M. (2011). La validit   du diagnostic issu d'un mariage entre didactique et mesure sur un test existant. Dans G. Ra  che, K. Paquette-C  t   et D. Magis (Eds), *Des m  canismes pour assurer la validit   de l'interpr  tation de la mesure en   ducation, volume 2* (pp. 11-30). Ste-Foy, Qu  bec: Presses de l'Universit   du Qu  bec.
- Maury, S. (2001). Didactique des math  matiques et psychologie cognitive : un regard comparatif sur trois approches psychologiques. *Revue fran  aise de p  dagogie*, 137, 85-93.

- 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 comp  tences    l'  cole : apprentissage et   valuation*, 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'  valuation des apprentissages dans une approche par comp  tences. 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'activit   et des comp  tences de professeurs des   coles et de professeurs de math  matiques dans des situations « simul  es » d'  valuation    vis  e formative en math  matiques. Th  se de doctorat non publi  e. Universit   Paris V.
- Vantourout, M. & Maury, S. (2006). Quelques r  sultats relatifs aux connaissances disciplinaires de professeurs stagiaires dans des situations simul  es d'  valuation de productions d'  l  ves en math  matiques. *Revue des sciences de l'  ducation*, vol. 32/3, 759-782.
- Vantourout, M. (2007).   tude de l'activit     valuative de professeurs stagiaires confront  s    des productions d'  l  ves en math  matiques : quel r  f  rent pour l'  valuateur ? *Mesure et   valuation en   ducation*, 30(3), 29-58.
- Vergnaud, G. (1983). Introduction. *Recherches en didactique des math  matiques*, Vol. 4/1, 9-25.
- Vergnaud, G. (1990). La th  orie des champs conceptuels. *Recherches en didactique des math  matiques*, Vol. 10/2.3, 133-170.
- Vergnaud, G. (2001). Psychologie du d  veloppement cognitif et   valuation des comp  tences. In Figari, G & Achouche, M. (Eds.), *L'activit     valuative r  interrog  e*, (pp. 43-51). Bruxelles : De Boeck.