

## Hierarchiser les besoins de diagnostic en mathématique en FP à l'aide d'un modèle de Rasch

Nathalie Loye and Carla Barroso da Costa

Volume 36, Number 2, 2013

URI: <https://id.erudit.org/iderudit/1024415ar>

DOI: <https://doi.org/10.7202/1024415ar>

[See table of contents](#)

Publisher(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (print)

2368-2000 (digital)

[Explore this journal](#)

Cite this article

Loye, N. & da Costa, C. B. (2013). Hierarchiser les besoins de diagnostic en mathématique en FP à l'aide d'un modèle de Rasch. *Mesure et évaluation en éducation*, 36(2), 59–85. <https://doi.org/10.7202/1024415ar>

Article abstract

According to the MELS (2008, 2010), 30% of students enrolled in a vocational training program (FP) in Quebec drop out or fail in their program of study. This is confirmed by principals and teachers of vocational training, especially when the program includes mathematical content, such as construction. This study therefore aims to better define the nature of difficulties in mathematics for students in vocational training. Using Rasch modeling, the study prioritizes diagnostic purposes in mathematics as perceived by teachers and school departments who completed a questionnaire. The results highlight the difficulties of students in several areas of mathematics and diagnostic purposes in particular with respect to logical reasoning, calculation techniques, approaches requiring multiple steps and understanding of the problem statement.

## Hiérarchiser les besoins de diagnostic en mathématique en FP à l'aide d'un modèle de Rasch

Nathalie Loye

Carla Barroso da Costa

Université de Montréal

**MOTS CLÉS:** évaluation diagnostique, attributs, mathématique, modèle de Rasch

*Selon le MELS (2008, 2010), 30% des élèves inscrits dans un programme de formation professionnelle (FP) au Québec n'obtiennent pas leur diplôme. Cette situation est confirmée par les directeurs et les enseignants de la formation professionnelle, particulièrement lorsque le programme inclut des contenus mathématiques comme, par exemple, en construction. La présente étude a donc pour objectif de mieux définir la nature des difficultés en mathématique des élèves de la formation professionnelle. En utilisant la modélisation de Rasch, l'étude hiérarchise les besoins de diagnostic en mathématique tels que perçus par les enseignants et directions des établissements scolaires qui ont rempli un questionnaire. Les résultats mettent en évidence les difficultés des élèves dans plusieurs domaines de la mathématique ainsi que les besoins de diagnostic, notamment en ce qui a trait au raisonnement logique, aux techniques de calcul, aux démarches nécessitant plusieurs étapes et à la compréhension de l'énoncé des problèmes.*

**KEY WORDS:** diagnostic assessment, attributes, mathematics, Rasch model

*According to the MELS (2008, 2010), 30% of students enrolled in a vocational training program (FP) in Quebec drop out or fail in their program of study. This is confirmed by principals and teachers of vocational training, especially when the program includes mathematical content, such as construction. This study therefore aims to better define the nature of difficulties in mathematics for students in vocational training. Using Rasch modeling, the study prioritizes diagnostic purposes in mathematics as perceived by teachers and school departments who completed a questionnaire. The results highlight the difficulties of students in several areas of mathematics and diagnostic purposes in particular with respect to logical reasoning, calculation techniques, approaches requiring multiple steps and understanding of the problem statement.*

PALAVRAS-CHAVE: avaliação diagnóstica, atributos, matemática, modelos de Rasch

*Segundo o MELS (2008, 2010), 30% dos alunos inscritos em um programa de formação profissional (FP) do Quebec evadem ou são reprovados nos estudos. Esta situação é confirmada pelos diretores e pelos professores da formação profissional, especialmente em situações em que o programa inclui conteúdos de matemática, como é o caso da área de construção civil. O presente estudo visa a melhor definir a natureza das dificuldades em matemática de alunos da formação profissional. À partir da utilização da modelização de Rasch, o estudo hierarquiza as necessidades de diagnóstico em matemática, tal como percebido pelos professores e diretores de estabelecimentos escolares que preencheram o questionário. Os resultados evidenciam as dificuldades dos alunos em diversas áreas da matemática, assim como necessidades de diagnóstico, principalmente no que tange ao raciocínio lógico, às técnicas de cálculo, aos procedimentos matemáticos compostos por várias etapas e à compreensão do enunciado dos problemas.*

## Introduction

Depuis plusieurs années, le ministère de l'Éducation du Loisir et du Sport (MELS) du Québec vise la réussite du plus grand nombre d'élèves, tant dans les programmes de la formation générale des jeunes et des adultes que dans ceux de la formation professionnelle. Actuellement, les programmes menant à l'obtention d'un diplôme d'enseignement professionnel (DEP) annoncent un taux moyen de diplomation d'environ 70 % (MELS, 2008, 2010). Toutefois, plusieurs discussions avec des directeurs et des enseignants des programmes de formation professionnelle (FP) font émerger leurs préoccupations face aux taux d'abandon et d'échec qu'ils jugent particulièrement importants dans les programmes incluant des contenus mathématiques, tels par exemple les programmes en construction dans lesquels la géométrie ou les calculs basés sur les nombres rationnels et les proportions sont essentiels.

En effet, la mathématique joue un rôle important dans plusieurs programmes de FP qui nécessitent le recours à des savoirs et à des compétences acquis dans cette discipline (MELS, 2007). De plus, l'anxiété due à la mathématique est pointée comme une cause de faible performance, et ce, dès le début du primaire (Vukovic, Kieffer, Bailey, & Harari, 2013). Certains auteurs vont jusqu'à parler de mathophobie (Gattuso, Lacasse, Lemire, & Van der Maren, 1989), la définissant comme un « état de panique, de paralysie et de désorganisation mentale [...] devant un problème de mathématiques » (p. 194). Cette mathophobie inclut à la fois des aspects sociaux, affectifs et cognitifs (Tobias, 1993).

Étant donné que la clientèle des centres de FP est de plus en plus hétérogène, incluant de plus en plus d'élèves ayant des besoins particuliers ou qui manquent de préalables (CRFP, 2009), cette étude est une première étape afin de tenter de mettre en évidence les aspects cognitifs des élèves de FP qui sont des entraves à leur réussite lorsque vient le temps d'utiliser des contenus et des compétences mathématiques dans leur formation.

Il s'agit de définir la nature des difficultés et de réfléchir à une manière de les mettre en évidence individuellement chez les élèves, dans une démarche d'évaluation diagnostique. L'hypothèse est que l'identification de ces difficultés pourrait permettre une remédiation efficace, avec comme corollaires une baisse de l'anxiété des élèves à l'égard des mathématiques et, à terme, un meilleur taux de réussite.

Comme point de départ pour cette première étape, il a été choisi d'interroger les enseignants et les directeurs de centres de FP afin de connaître leur perception des difficultés de leurs élèves en mathématique. Cet article rapporte donc les résultats d'une enquête menée entre mai et septembre 2011 dans les centres de FP de la grande région de Montréal. L'application d'un modèle de Rasch aux données vise à hiérarchiser les besoins de diagnostic en mathématique tels que perçus par les enseignants ou par le personnel de direction, posant ainsi les bases de la démarche qui mènera ensuite à élaborer des tests diagnostiques en mathématique.

### **La visée diagnostique de l'évaluation**

Traditionnellement, l'évaluation vise d'abord la certification et la diplomation. Avec le renouveau pédagogique, les programmes de formation actuels mettent une emphase particulière sur la visée formative de l'évaluation qui consiste à utiliser l'évaluation en salle de classe pour aider les élèves à apprendre. Cette évaluation pour l'apprentissage repose sur une rétroaction en salle de classe, réalisée de manière instrumentée ou pas, et a pour objectif de guider l'élève dans ses apprentissages (MELS, 2003). La visée diagnostique s'inscrit, pour certains auteurs, dans cette démarche d'évaluation formative à la grandeur de la salle de classe. Pour une discussion à ce sujet, voir Marcoux, Fagnant, Loye, et Ndinga (sous presse).

Pour d'autres auteurs (par ex., Leighton & Gierl, 2007; Loye, 2010; Rupp, Templin, & Henson, 2010), la visée diagnostique dépasse la dimension de la salle de classe et doit s'appuyer sur des tests ayant le pouvoir de mettre en évidence les attributs qui posent problème à chaque élève à partir d'une liste préétablie. Le diagnostic repose alors sur une modélisation psychométrique des réponses fournies par les élèves aux items du test (voir par ex., Loye, 2010) et doit bien sûr être suivi d'un plan de remédiation adapté à chaque élève.

La présente étude s'inscrit dans cette deuxième perspective et vise, à terme, à élaborer des tests qui pourraient être utilisés à grande échelle auprès des élèves de FP, pour mettre en évidence individuellement les attributs qui posent problème et ceux qui n'en posent pas. Dans une telle démarche, l'enseignant pourrait, par exemple, recevoir le rapport diagnostique individuel de chacun de ses élèves et un guide pour les aider à surmonter les difficultés rele-

vées. Dans la perspective d'élaborer de tels tests diagnostiques, la première étape consiste donc à définir clairement quels devraient être les attributs à diagnostiquer (Loye et al., 2011).

### **Les attributs à diagnostiquer selon la littérature**

Un attribut fait référence à tout ce qui peut affecter la performance d'un élève lors de la résolution d'une tâche et peut être une caractéristique de la tâche ou du contexte, ou faire référence aux microniveaux d'habiletés nécessaires pour trouver la bonne réponse (Gierl, 1997 ; Tatsuoka, 2009). Il peut s'agir d'un savoir, d'un savoir-faire, d'une stratégie qui ne peuvent pas être observés directement, mais qui peuvent être combinés pour fournir la bonne réponse aux items du test (Gierl, 1997).

Dans le domaine de la mathématique, de nombreux auteurs ont proposé des listes d'attributs (Birenbaum, Tatsuoka, & Yamada, 2004 ; Hartz, 2002 ; Milewski & Baron, 2002). Dans une précédente étude (Loye, 2008), quatre experts ont défini une liste d'attributs en lien avec quarante items issus des épreuves diagnostiques en mathématique de l'école Polytechnique de Montréal. Ces experts, qui étaient des enseignants avec une formation en didactique des mathématiques, ont travaillé individuellement à déterminer ces attributs. De la liste de 106 attributs qui a résulté, un classement consensuel obtenu par la méthode *Multi Attribute Consensus Building* (MACB) (Vanderwood, Ysseldyke, & Thurlow, 1993) a permis de faire des regroupements en six catégories :

- 1) décoder : tout ce qui a trait au fait de lire la question correctement,
- 2) appliquer une technique : calcul algébrique, arithmétique, application de règle ou d'algorithme,
- 3) faire des liens entre des notions,
- 4) utiliser une visualisation graphique,
- 5) utiliser des propriétés : définitions, théorèmes, et
- 6) organiser la solution.

Ces catégories permettent de caractériser la plupart des attributs proposés par les différents auteurs cités plus haut, ce qui amène à les utiliser pour la présente étude.

## Méthodologie

Entre mai et septembre 2011, un sondage Web a été mené auprès des enseignants et du personnel de direction de plusieurs centres de FP de la grande région de Montréal, afin de mettre à profit leur expérience pour cibler et hiérarchiser les attributs utiles à diagnostiquer.

### *Les participants*

Le sondage a été envoyé par courriel à une trentaine de centres de FP de la grande région de Montréal, choisis en fonction de leur offre de programmes incluant des mathématiques et de l'acceptation de leur directeur à diffuser le sondage parmi son équipe. Au total, 51 personnes ont rempli le sondage. Les répondants sont en grande majorité ( $n = 40$ ) des enseignants, certains enseignent la mathématique ( $n = 3$ ), les autres utilisent des contenus mathématiques dans leur enseignement ( $n = 37$ ). Sept membres de direction, deux conseillers en orientation, un conseiller pédagogique et un conseiller en information scolaire ont également répondu au sondage.

### *Le questionnaire*

Le questionnaire comportait trois sections. La première section visait à connaître la fonction exercée par le répondant, le domaine de formation et le contexte dans lequel il intervient. La seconde section portait sur sa perception générale quant aux taux de réussite des élèves. Elle incluait deux questions ouvertes, l'une sur les raisons perçues des difficultés des élèves en mathématique et l'autre sur les approches pédagogiques utilisées pour pallier ces difficultés.

Enfin, la troisième section portait sur la perception du répondant relativement au niveau de difficulté des élèves. Un premier ensemble d'items (Q13) ciblait des contenus mathématiques enseignés en FP :

- 1) concepts et opérations de base sur les fractions,
- 2) concepts et opérations de base sur les proportions,
- 3) concepts et opérations de base en algèbre, et
- 4) concepts et opérations de base en géométrie plane.

L'échelle proposée comportait huit niveaux de 1 – *très facile* à 8 – *très difficile* pour les élèves.

Un deuxième ensemble d'items (Q14) portait sur l'utilité perçue d'avoir accès à un diagnostic relatif à 14 attributs tirés de la liste initiale de Loye (2008) et retenus pour caractériser au mieux, dans le contexte de la FP, chacune des six catégories énoncées plus haut. La liste initiale ayant été conçue par des experts qui étaient enseignants, le recours à un seul enseignant a été jugé utile dans le cadre de la présente étude pour valider le choix de ces 14 attributs. Notons que ces attributs font uniquement référence au processus et ne tiennent pas compte des contenus spécifiques. Ces attributs et leur classification dans les six catégories font l'objet du tableau 1.

Tableau 1  
*Liste des attributs en mathématique retenus pour chaque catégorie*

<i>Catégories</i>	<i>Attributs</i>
1) <b>Décoder</b>	- Comprendre des problèmes posés sous forme de phrases - Comprendre des problèmes posés en utilisant le langage mathématique - Reconnaître les mots clés dans l'énoncé - Lire et comprendre (plus petit que, doit, devrait, au moins, négation, augmenter, diminuer, etc.)
2) <b>Appliquer une technique</b>	- Appliquer une technique de calcul - Calculer (calculs du genre à être facilités par une calculatrice) - Manipuler une expression algébrique
3) <b>Faire des liens entre des notions</b>	- Reasonner logiquement - Évaluer et vérifier des options
4) <b>Utiliser une visualisation graphique</b>	- Utiliser des figures, tableaux ou graphiques
5) <b>Utiliser des propriétés</b>	- Utiliser des propriétés, définitions, etc.
6) <b>Organiser la solution</b>	- Traduire des mots en expression mathématique - Appliquer une démarche ayant une seule étape - Appliquer une démarche ayant plusieurs étapes

L'échelle proposée comportait huit niveaux de 1 – *peu utile* à 8 – *très utile* pour l'enseignant. Dans les deux cas, un choix «je ne sais pas» était proposé. Une question ouverte demandant de décrire ses besoins en matière de test diagnostique clôturait le sondage.



Les formulations des questions du sondage ont été validées par quatre enseignantes en mathématique avant de procéder à la collecte des données à l'aide d'un sondage Web.

### ***Les analyses***

Dans un premier temps, une analyse descriptive des données permet de faire un portrait rapide des classes de FP auxquelles les répondants font référence. Les deux questions de la section 3 sont ensuite analysées en deux étapes. La première est une étude de la dimensionnalité des items qui composent Q13 ( $\alpha = 0,90$ ) et Q14 ( $\alpha = 0,83$ ) à l'aide d'une analyse factorielle. Par la suite, les données de ces deux questions sont analysées à l'aide d'un modèle de Rasch en utilisant le logiciel RUMM2020 (Andrich, Lyne, Sheridan, & Luo, 2003) dans le but de hiérarchiser les besoins de diagnostic en ce qui a trait aux contenus mathématiques, d'une part (Q13), et aux 14 attributs ciblés, d'autre part (Q14). Enfin les réponses à la dernière question ouverte du questionnaire sont présentées.

## **Résultats**

### ***La description des classes de FP***

Les données récoltées dans les deux premières sections du questionnaire permettent de brosser un portrait des classes de FP. À une grande majorité, celles-ci sont qualifiées d'hétérogènes et incluent entre 16 et 25 élèves. Pour 74 % des répondants, elles comportent moins de 20 % d'élèves provenant de communautés culturelles variées. En ce qui concerne la mathématique, 93 % des répondants considèrent que les élèves sont au mieux de niveau moyen, et plus d'un quart d'entre eux les considèrent faibles ou très faibles en mathématique. Le taux de réussite dans les cours impliquant des contenus en mathématique est toutefois supérieur à 50 % pour la quasi-totalité des répondants, la moitié d'entre eux le considère supérieur à 75 %. Finalement, 80 % des répondants estiment que le taux de réussite dans leur programme est supérieur à 75 %.

Le tableau 2 présente la synthèse des raisons des difficultés des élèves invoquées par les répondants, et le tableau 3 récapitule les approches pédagogiques proposées pour pallier ces difficultés.

Tableau 2  
*Raisons des difficultés des élèves selon les répondants*

<i>Raisons</i>	<i>Nombre de répondants</i>
Manque ou fragilité des notions de base	16
Calcul mental déficient ou habitude d'utiliser la calculatrice	6
Manque de logique, de compréhension, de rigueur	5
Fait que les élèves retournent à l'école après un arrêt dans leurs études	5
Secondaire qui ne les prépare pas adéquatement	4
Poids des difficultés passées	2
Manque de pratique	2
Manque d'intérêt	1

Tableau 3  
*Approches pédagogiques proposées par les répondants*

<i>Approches pédagogique</i>	<i>Nombre de répondants</i>
Exercices supplémentaires, récupération	21
Proposer des exercices concrets en lien avec le métier	8
Faire un test diagnostique	2
Ne sais pas	2
Ajouter un enseignant pour aider à faire les exercices	1
Varier les exercices	1

La principale raison des difficultés des élèves est donc le manque ou la fragilité des notions de base pour lesquelles l'approche pédagogique préconisée est la récupération. Plusieurs participants ajoutent que l'attention doit être portée à proposer des problèmes en lien avec le métier. Enfin, un commentaire cible le besoin de varier les types d'exercices et de sortir des habituelles manières de faire.

### *Les résultats liés à la section 3 du questionnaire*

#### *Répartition des répondants pour Q13 et Q14*

Des 51 personnes qui ont rempli le questionnaire, seules 35 ont fourni des réponses à tous les items de la question Q13, et 34 à tous ceux de la question Q14. Le tableau 4 présente la répartition de ces répondants selon leur fonction et le domaine de formation dans lequel ils interviennent. Il est possible de vérifier que la majorité d'entre eux ( $n = 24$ ) sont des enseignants qui utilisent des contenus mathématiques dans plusieurs domaines.

Tableau 4  
*Répartition des répondants aux questions Q13 et Q14*

<i>Programmes de FP</i>									
	Sécurité et incendie	Construction	Aérospatiale	Mécanique d'entretien	Métallurgie	Transports	Alimentation	Autres	Total
Enseignant en mathématique	0	1	0	0	0	0	0	2	3
Enseignant qui utilise des contenus mathématiques	5	9	0	2	2	0	1	5	24
Conseiller en orientation	0	0	1	0	0	1	0	0	2
Personnel de direction	1	1	0	0	2	0	1	1	6
Total	6	11	1	2	4	1	2	8	35

#### *Étude de la dimensionnalité des questions Q13 et Q14*

La question Q13 est constituée de quatre items portant sur la difficulté perçue par les répondants de quatre contenus mathématiques. La question Q14 comporte 14 items sur l'utilité perçue d'avoir accès à un diagnostic relativement à une liste de 14 attributs. Tous les items sont polytomiques sur une échelle à huit points, les réponses «je ne sais pas» ont été considérées comme des valeurs manquantes. Dans ce cas, les valeurs manquantes sont considérées

comme étant manquantes au hasard (*Missing At Random* - MAR) (elles dépendent ici en partie du statut du répondant), et peuvent souvent être ignorées (Shafer & Graham, 2002). Elles constituent moins de 10 % des données pour les deux questions.

Les données recueillies auprès des répondants ont d'abord fait l'objet d'analyses factorielles séparées afin d'étudier la dimensionnalité de Q13 et de Q14. La taille de l'échantillon ( $n = 35$  ou  $n = 34$ ) est plus petite que 100, valeur indiquée comme seuil minimal pour les analyses factorielles par plusieurs auteurs (par ex., Field, 2009 ; Tabachnick & Fidell, 2007). Toutefois, les auteures se sont appuyées sur des études qui montrent la possibilité de les réaliser avec de petits échantillon ( $n < 50$ ) lorsque les facteurs sont bien définis, ou limités à une ou deux dimensions, et que les valeurs de communauté sont plutôt élevées (Mundfrom, Shaw, & Ke, 2005 ; Winter, Dodou, & Wieringa, 2009). Les auteures ont procédé aux analyses à l'aide de SPSS en utilisant la méthode d'extraction « factorisation par axes principaux », la rotation orthogonale « varimax » et une saturation plus élevée que 0,30.

Pour Q13, le test de sphéricité de Bartlett ( $\chi^2(6) = 30,719, p = 0,000$ ) ainsi que la mesure Kaiser-Meyer-Olkin ont indiqué un ajustement acceptable des items aux facteurs latents (KMO = 0,699). La solution à un facteur expliquait 55,40 % de la variance totale, en conservant les quatre items. Tous les items ont présenté une saturation supérieure à 0,40, dépassant le seuil imposé de 0,30, tel que présenté dans le tableau 5.

Tableau 5  
*Indices de saturation pour les items de Q13*

	<i>Items</i>	<i>Saturation</i>
Q1301	Concepts et opérations de base sur des fractions	0,582
Q1302	Concepts et opérations de base dans des proportions	0,838
Q1303	Concepts et opérations de base en algèbre	0,717
Q1304	Concepts et opérations de base en géométrie 2D	0,406

Pour Q14, le test de sphéricité de Bartlett ( $\chi^2(91) = 326,096, p = 0,000$ ) ainsi que la mesure Kaiser-Meyer-Olkin ont indiqué un bon ajustement des items aux facteurs latents (KMO = 0,730). Comme pour Q13, la solution à un facteur expliquait 50,84 % de la variance totale, en conservant les 14 items. Néanmoins, la saturation de l'item q1401 (*utiliser des figures, tableaux ou*

*graphiques*) était inférieure aux autres (0,366), suggérant la possibilité de rencontrer des problèmes d'ajustement lors des analyses de Rasch. Le tableau 6 présente l'ensemble des indices de saturation pour la question Q14.

Tableau 6  
*Indices de saturation pour les items de Q14*

	<i>Items</i>	<i>Saturation</i>
Q1401	Utiliser des figures, tableaux ou graphiques	0,366
Q1402	Comprendre des problèmes posés sous forme de phrases	0,721
Q1403	Comprendre des problèmes posés en utilisant le langage mathématique	0,677
Q1404	Reconnaître les mots clés dans l'énoncé	0,856
Q1405	Traduire des mots en expressions mathématiques	0,725
Q1406	Utiliser des propriétés, définitions, etc.	0,690
Q1407	Appliquer une technique de calcul	0,847
Q1408	Manipuler une expression algébrique	0,545
Q1409	Calculer (calculs du genre à être facilités par une calculatrice)	0,596
Q1410	Lire et comprendre (plus petit que, doit, devrait, au moins, négation, augmenter, diminuer...)	0,639
Q1411	Raisonnement logiquement	0,663
Q1412	Appliquer une démarche ayant une seule étape	0,591
Q1413	Appliquer une démarche ayant plusieurs étapes	0,752
Q1414	Évaluer et vérifier des options	0,816

Ces analyses préliminaires ont montré que les deux questions Q13 et Q14 étaient raisonnablement unidimensionnelles et ensuite les données sont modélisées à l'aide du modèle de Rasch.

### *Modélisations de Rasch*

Les données de Q13 et Q14 ont été analysées séparément à l'aide d'un modèle de Rasch à crédit partiel qui permet d'adapter l'échelle à chacune des questions (Bond & Fox, 2007), et ainsi de tenir compte d'une éventuelle inconsistance dans l'utilisation des échelons par les répondants. L'avantage est l'obtention d'un modèle mieux adapté aux données.

Comme pour les analyses factorielles, la pertinence d'une telle modélisation avec un petit nombre de répondants doit être remise en question. Tout d'abord, plusieurs auteurs justifient la pertinence d'appliquer une modélisation de Rasch à des échantillons de petite taille (Linacre, 1994; Lord, 1983). Selon Andrich (2013), même avec un modèle à crédit partiel, le fait qu'une étude vise essentiellement à porter un regard général sur des données dans une perspective exploratoire est un argument en faveur de l'approche choisie, mais nécessite de nuancer les interprétations des résultats. Il faut, par exemple, être attentif à l'adéquation des distributions des répondants et des items sur l'échelle de mesure et à la taille de l'erreur standard des paramètres estimés. De plus, les résultats donnent un ordre d'idée de la position sur l'échelle de mesure et non une mesure précise.

Pour les quatre items de Q13, l'ajustement du modèle initial est acceptable avec une interaction item-trait non significative ( $\chi^2 = 2,277$ ;  $df = 8$ ;  $p = 0,971$ ). La moyenne des résidus pour les items est de 0,267 avec un écart type de 0,948, celle pour les répondants est de -0,417 avec un écart type de 1,093 (voir tableau 7).

Pour les 14 items de Q14, l'ajustement du modèle initial met en évidence un problème d'ajustement avec une interaction item-trait statistiquement significative ( $\chi^2 = 58,60$ ;  $df = 28$ ;  $p = 0,000$ ). Les moyennes des résidus pour les items et pour les répondants sont respectivement 0,557 avec un écart type de 1,090 et 0,032 avec un écart type de 1,384. Autant pour Q13 que pour Q14, les indices de séparation (PSI) présentent une consistance interne supérieure à 0,70, considérée acceptable selon Ramp, Khan, Mijason, et Pallant (2009), et adéquate en fonction de la quantité réduite d'items (Pallant & Tennant, 2007). Le tableau 7 présente les résultats des itérations qui ont permis d'ajuster les données au modèle et de choisir le modèle final pour Q13 et Q14.

Tableau 7  
*Ajustements itératifs des modèles pour Q13 et Q14*

<i>Description du modèle</i>	<i>Ajustement global Moy. (écart type)</i>	<i>Ajustement item Moy. (écart type)</i>	<i>Ajustement personne</i>	<i>PSI</i>
<b>Q 13</b>				
(1) Modèle initial	$\chi^2 = 2,277$ df = 8 $p = 0,971$	0,267 (0,948)	-0,417 (1,093)	0,837
(2) Modèle après changement d'échelle q1304	$\chi^2 = 6,458$ df = 8 $p = 0,596$	0,219 (0,856)	-0,372 (1,054)	0,827
(3) Modèle après avoir retiré le cas extrême	$\chi^2 = 6,695$ df = 8 $p = 0,570$	0,185 (0,799)	-0,284 (0,843)	0,820
(4) Modèle après nouveau changement d'échelle q1304 *	$\chi^2 = 6,202$ df = 8 $p = 0,625$	0,208 (0,731)	-0,320 (0,894)	0,825
<b>Q 14</b>				
(1) Modèle initial	$\chi^2 = 58,60$ df = 28 $p = 0,000$	0,557 (1,090)	0,032 (1,384)	0,946
(2) Modèle après changement d'échelle pour 10 items	$\chi^2 = 38,316$ df = 28 $p = 0,093$	0,434 (1,067)	-0,076 (1,326)	0,904
(3) Modèle sans item q1401	$\chi^2 = 34,203$ df = 26 $p = 0,129$	0,446 (0,920)	-0,075 (1,260)	0,902
(4) Modèle sans le cas extrême	$\chi^2 = 34,149$ df = 26 $p = 0,131$	0,444 (0,881)	-0,014 (1,118)	0,897
(5) Modèle après nouveau changement d'échelle q1403	$\chi^2 = 29,078$ df = 26 $p = 0,308$	0,421 (0,886)	-0,069 (1,208)	0,897
(6) Modèle après avoir retiré un sujet*	$\chi^2 = 24,986$ df = 26 $p = 0,519$	0,436 (0,881)	0,011 (1,113)	0,894

Note. \* Modèle final

La démarche itérative pour ajuster les données au modèle consiste tout d'abord à examiner l'ordre des échelons pour chacun des items. En effet, un désordre observé indique un manque de cohérence de la part des participants dans l'utilisation des échelons proposés dans le questionnaire. En ce qui concerne Q13, seul l'item q1304 ne respecte pas l'ordre des échelons, tel qu'illustré par la figure 1, puisque les pas 1, 2 et 6 n'ont jamais une probabilité supérieure aux autres. En ce qui concerne Q14, dix items présentent des problèmes similaires dans le modèle initial sans qu'il soit toutefois possible d'en discerner la cause. Les itérations successives des modélisations tiennent compte de ce problème et visent à définir le nombre de catégories correspondant à chaque item.

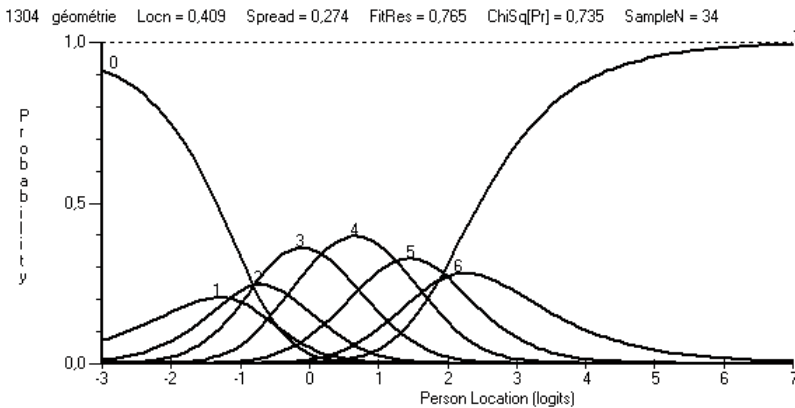


Figure 1. *Exemple pour illustrer des problèmes liés à l'ordre des échelons (item q1304)*

La démarche consiste ensuite à identifier les répondants et les items qui présentent un problème d'ajustement identifié lorsque le résidu correspondant est inférieur à  $-2,5$  ou supérieur à  $+2,5$  et à les retirer des analyses.

Selon la démarche modélisée par Pallant et Tennant (2007), quatre itérations ont été nécessaires pour ajuster le modèle à crédit partiel pour Q13 (voir tableau 7). Après avoir ajusté l'échelle en huit points de l'item q1304 (0,1,2,3,4,5,6,7) en une échelle en trois points (0,1,1,1,2,2,2,2) et éliminé un sujet, le modèle final a un bon ajustement ( $\chi^2 = 6,202$ ;  $df = 8$ ;  $p = 0,625$ ) avec un indice de séparation (PSI) de 0,825. Les indices d'ajustement des items ont une moyenne de 0,208 avec un écart type de 0,731 ; ceux des répondants ont une moyenne de  $-0,320$  et un écart type de 0,894. Ces valeurs sont relativement proches des valeurs espérées de 0 et 1. Le tableau 8 présente les





Tableau 9  
*Ajustement des items au modèle final pour Q14*

<i>Item</i>	<i>Logit</i>	<i>SE</i>	<i>Résidu</i>	<i>DL</i>	<i>Chi<sup>2</sup></i>	<i>DL</i>	<i>Prob.</i>
1402 Comprendre des problèmes posés sous forme de phrases	0,232	0,173	0,971	23,17	5,118	2	0,077
1403 Comprendre des problèmes posés en utilisant le langage mathématique	0,387	0,357	0,128	22,37	0,927	2	0,629
1404 Reconnaître les mots clés dans l'énoncé	0,663	0,179	-0,646	23,17	0,490	2	0,783
1405 Traduire des mots en expression mathématique	1,161	0,243	0,793	21,57	2,292	2	0,318
1406 Utiliser des propriétés, définitions, etc.	1,435	0,161	0,931	20,77	1,167	2	0,558
1407 Appliquer une technique de calcul	-1,846	0,350	-0,420	23,97	1,277	2	0,529
1408 Manipuler une expression algébrique	0,859	0,319	<b>2,680</b>	20,77	2,170	2	0,338
1409 Calculer (calculs du genre à être facilités par une calculatrice)	0,446	0,289	0,298	22,37	1,983	2	0,371
1410 Lire et comprendre (plus petit que, doit, devrait, au moins, négation, augmenter, diminuer, etc.)	-1,522	0,348	0,781	23,97	0,844	2	0,656
1411 Raisonner logiquement	-2,117	0,360	-0,395	23,97	0,167	2	0,920
1412 Appliquer une démarche ayant une seule étape	0,768	0,171	0,591	21,57	3,477	2	0,176
1413 Appliquer une démarche ayant plusieurs étapes	-1,068	0,310	-0,441	21,57	3,632	2	0,163
1414 Évaluer et vérifier des options	0,603	0,232	0,401	20,77	1,443	2	0,487

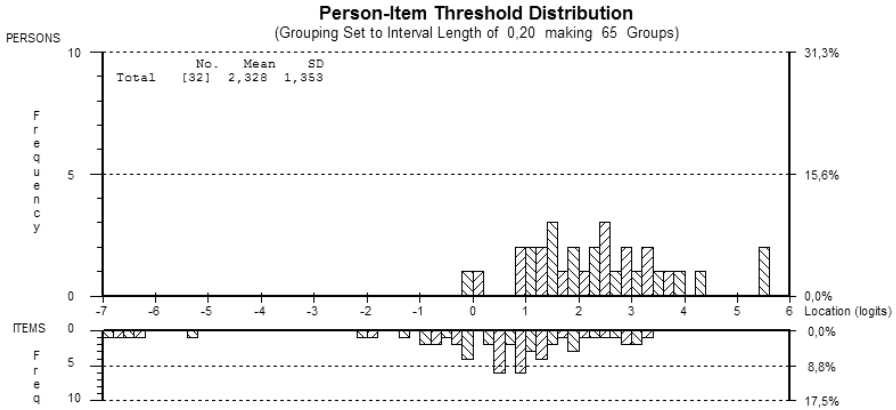


Figure 3. *Adéquation des items et des sujets pour Q14*

En outre, l'application d'un modèle de Rasch repose sur un postulat d'indépendance locale en lien avec l'unidimensionnalité des données. Pallant et Tenant (2007) suggèrent de faire une analyse en composantes principales (ACP) des résidus afin de détecter une éventuelle violation de ce postulat. Concrètement, la démarche propose de réaliser une série de tests t pour échantillons appariés afin de vérifier si les *logits* des répondants diffèrent de manière significative entre leurs scores liés aux items ayant une valeur positive et ceux liés aux items ayant une valeur négative dans le premier facteur de l'ACP. Pour que l'échelle soit considérée unidimensionnelle, au plus 5% des cas doivent avoir une différence significative (avec un intervalle de confiance de 95%) entre les scores aux deux ensembles d'items. Les résultats indiquent, autant pour Q13 que pour Q14, que tous les tests t sont non significatifs. Ceci permet de soutenir le postulat d'indépendance locale et d'unidimensionnalité en cohérence avec les résultats des analyses factorielles réalisées préalablement.

Finalement, les figures 4 et 5 présentent les positions (*logits*) des répondants et des items sur l'échelle de mesure respectivement pour Q13 et Q14.

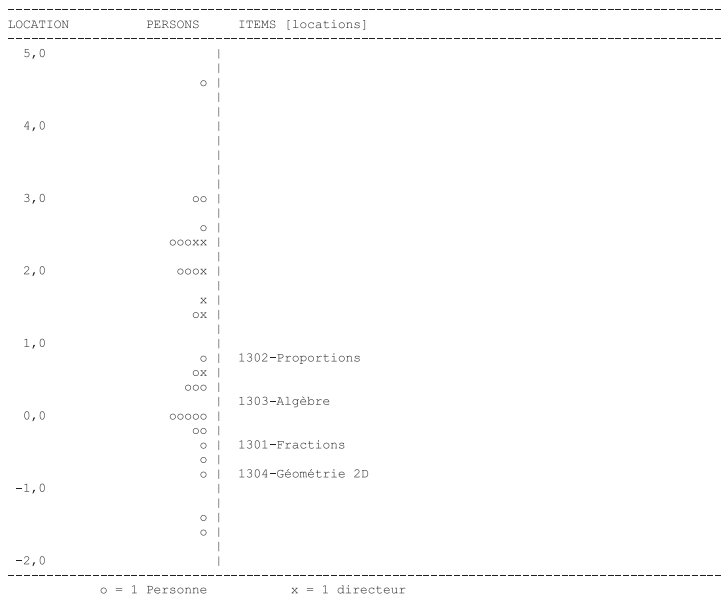


Figure 4. *Position des sujets et des items sur l'échelle de mesure pour Q13*

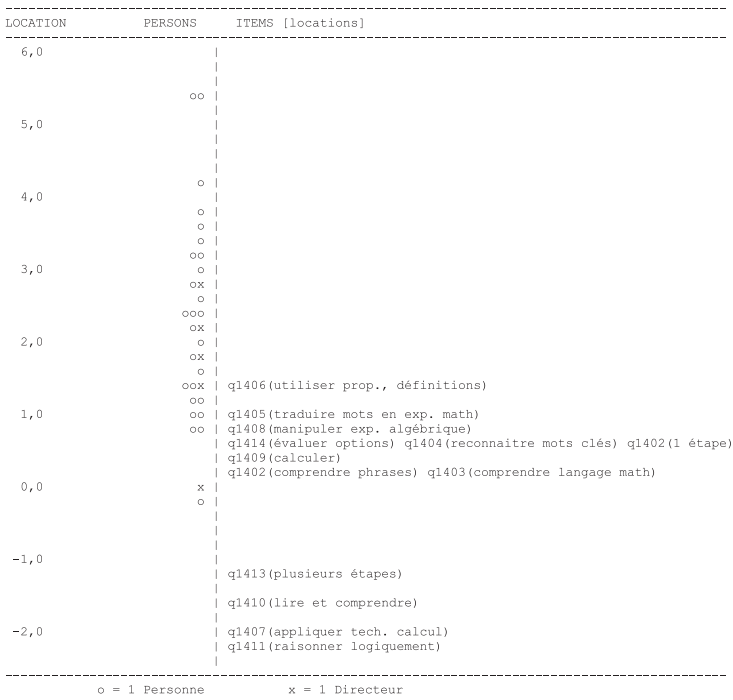


Figure 5. *Position des sujets et des items sur l'échelle de mesure pour Q14*

Étant donné le petit nombre de répondants, les erreurs standards des positions des items sur l'échelle de mesure ont été utilisées pour mieux différencier les valeurs. La figure 6 met ainsi en évidence l'existence de deux groupes d'items pour Q13, de même que pour Q14.

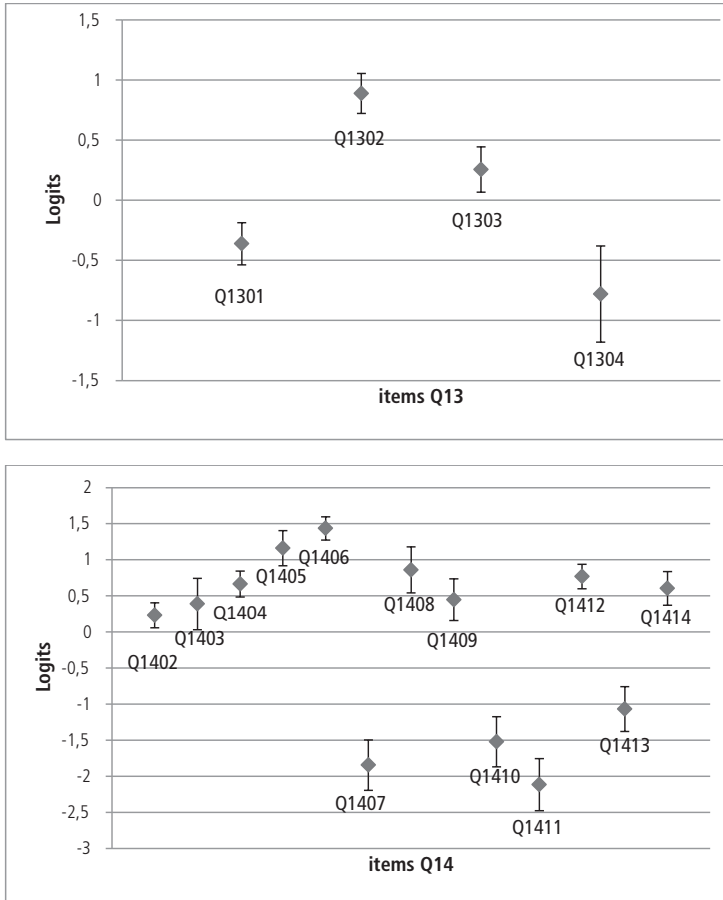


Figure 6. *Position des items sur l'échelle de mesure (logits) pour Q13 et Q14 incluant l'erreur standard (ES)*

### *Autres besoins exprimés*

La dernière question donnait la possibilité aux répondants d'exprimer librement leurs besoins en matière de test diagnostique. Les réponses (voir tableau 10) pointent vers la nécessité de prendre en compte des difficultés liées :

- 1) aux unités de mesure et notamment aux conversions entre les systèmes métrique et impérial,
- 2) aux calculs mental ou par écrit mais sans usage de calculatrice, et
- 3) au raisonnement logique.

De plus, quatre répondants insistent sur la nécessité de proposer une épreuve dans laquelle la mathématique est directement appliquée au domaine de formation de l'élève.

Tableau 10  
*Besoins de diagnostic mentionnés par les répondants*

<i>Besoins</i>	<i>Nombre de répondants</i>
Systèmes de mesure métrique et impérial	6
Notions de base en calcul	5
Mathématiques appliquées au métier	4
Logique, raisonnement	3
Géométrie	1
Bonne lecture des données	1
Dépister les lacunes, diagnostic général	3
Aucun besoin	2

### **Interprétation et discussion**

Cette étude constitue la première étape d'un programme de recherche visant à développer des tests diagnostiques en mathématique adaptés aux élèves de FP. Son objectif était de vérifier les besoins du milieu et de guider les premières étapes de ce développement. Le premier constat est que les taux de réussite rapportés par les répondants sont comparables à celui de 70% en moyenne annoncé par le MELS. Toutefois, les réponses mettent bien en évidence qu'une proportion importante d'élèves présente des difficultés en mathématique.

Les auteures souhaitaient établir une hiérarchie des difficultés et de l'utilité de divers diagnostics en mathématique selon les perceptions des enseignants et membres de la direction des centres de FP qui ont répondu au questionnaire. Les diverses données recueillies à l'aide du questionnaire ont permis de documenter les besoins. Malgré le nombre de répondants relativement petit, les résultats montrent que les données de Q13 et de Q14 ont pu être modélisées avec un modèle de Rasch à crédit partiel en présentant un bon ajustement.

Les items de Q13 sont en lien avec le niveau de difficulté de quatre contenus mathématiques. À la figure 4, l'échelle commune aux participants et aux items s'étend de  $-2 \text{ logits}$  à  $+5 \text{ logits}$ . Pour les participants, elle représente leur propension à trouver les items difficiles. Ainsi, plus la valeur de son *logit* est élevée, plus il est facile pour le participant de trouver difficiles les items placés plus bas que lui sur le continuum. En ce qui concerne les items, l'échelle représente leur niveau de difficulté. Plus la valeur du *logit* d'un item est petite, et plus l'item a tendance à être considéré comme difficile par les participants placés plus haut que lui.

Ainsi, l'interprétation de la figure 4 permet de constater que la géométrie et les fractions sont jugées plus difficiles que l'algèbre et les proportions. La figure 6, qui inclut les erreurs standards, confirme l'existence de deux groupes d'items. En outre, à la figure 4, les directeurs sont identifiés par les x. Le fait que les x se situent dans la partie supérieure de la figure semble mettre en évidence leur propension à trouver tous les contenus mathématiques difficiles pour les élèves, un certain nombre d'enseignants étant plus modérés qu'eux.

De manière générale, les participants de l'étude considèrent que les élèves ont des difficultés en mathématiques et les analyses permettent une certaine hiérarchisation de ces difficultés qui vient confirmer les impressions issues de discussions informelles avec des enseignants et des directeurs de centres de FP. Il faut donc particulièrement tenir compte des difficultés en lien avec les fractions et la géométrie dans l'élaboration des items des tests diagnostiques, en gardant à l'esprit que les notions de base des élèves sont fragiles.

Les items de Q14 sont en lien avec l'utilité d'avoir accès à une information diagnostique relativement à 14 attributs. À la figure 5, l'échelle commune aux participants et aux items s'étend également de  $-2 \text{ logits}$  à  $+5 \text{ logits}$ . Pour les participants, elle représente leur propension à trouver utile un diagnostic relativement aux attributs. Ainsi, plus la valeur de son *logit* est élevée, plus il est facile pour le participant de trouver utile un diagnostic lié aux items placés plus bas que lui sur le continuum. En ce qui concerne les items, l'échelle représente leur niveau d'utilité. Plus la valeur du *logit* d'un item est petite, et plus l'item a tendance à être considéré comme utile par les participants placés plus haut que lui.

L'interprétation de la figure 5 permet de constater que l'ensemble des 14 attributs est considéré comme utile à diagnostiquer par la quasi-totalité des répondants. Toutefois, il est possible de distinguer deux regroupements, confirmés par la figure 6 qui inclut les erreurs standards. Le premier est constitué de

quatre attributs qui se détachent clairement par le fait que la totalité des répondants sont placés plus hauts qu'eux sur le continuum. Ainsi, *raisonner logiquement*, *appliquer une technique de calcul* sont les attributs les plus importants à prendre en compte dans un test diagnostic en mathématique. Il est intéressant de constater que ces deux attributs se retrouvent de manière explicite, autant dans les raisons des difficultés que dans les besoins de diagnostics proposés par les répondants. Le fait de *lire et comprendre* et celui de *procéder en plusieurs étapes* pour répondre à la question complètent ce premier groupe d'attributs considérés comme étant les plus utiles à diagnostiquer.

Le deuxième regroupement est constitué par les neuf autres attributs qui sont considérés comme utiles à diagnostiquer par la quasi-totalité des répondants. Il est possible d'y retrouver *calculer (calculs du genre à être facilités par une calculatrice)*, *résoudre en une seule étape* et *manipuler des expressions algébriques*. Cinq attributs font référence à divers aspects liés à la compréhension de l'énoncé. Le besoin d'*utiliser des propriétés et des définitions* complète ce regroupement. Notons qu'*utiliser des figures, tableaux ou graphiques* a été éliminé dans le processus d'ajustement des données au modèle, peut-être parce que les élèves de FP le font peu.

Malgré le petit nombre de répondants, des analyses factorielles et une modélisation de Rasch ont été effectuées. Même si ces choix méthodologiques peuvent être discutables et constituent une limite à cette étude en raison notamment du possible manque de précision dans l'estimation des paramètres des items et des répondants, nous considérons que le caractère exploratoire de l'étude et les conclusions générales tirées sont en accord avec ces choix.



## Conclusion

Les résultats de cette étude ont permis de documenter, selon la perception des personnes qui ont rempli le questionnaire, les difficultés en mathématique vécues par les élèves en FP dans les quatre domaines ciblés par les programmes de formation et qui sont la géométrie, les fractions, l'algèbre et les proportions.

Ils pointent également vers le fait qu'une information diagnostique est perçue comme plus importante pour quatre des 14 attributs. Ainsi, les auteures travaillent actuellement à développer des tests diagnostiques visant à mettre en évidence des difficultés spécifiquement en lien avec le *raisonnement logique*, l'*application de techniques de calcul*, la *compréhension des énoncés* et la *résolution de problème nécessitant plusieurs étapes*. Les tests qui seront créés s'inscrivent de manière authentique dans le domaine de formation des candidats et il importe de veiller à couvrir les contenus spécifiques à chaque programme dans des tests adaptés à chacun d'eux.

## RÉFÉRENCES

- Andrich, D. (2013). Forum de discussion MBC-Rasch- *sample size*. Consulté à l'adresse <https://mailman.wu.ac.at/mailman/private/rasch//2013q2/003444.html>
- Andrich, D., Lyne, A., Sheridan, B., & Luo, G. (2003). RUMM 2020 [Computer Software]. Perth, Australia: RUMM Laboratory.
- 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. doi: <http://dx.doi.org/10.1016/j.stueduc.2004.06.004>
- Bond, T., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences*. (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- CRFP (2009). *La réussite scolaire en formation professionnelle*. Journée d'étude du 13 février 2008. Extrait de <http://pages.usherbrooke.ca/crfp/pdf/rapport%20journee%20d%27etude.pdf>
- Field, A. (2009). *Discovering Statistics Using SPSS*. London, UK: Sage publication.
- Gattuso, L., Lacasse, R., Lemire, V., & Van der Maren, J. (1989). Quelques aspects sociaux et affectifs de l'enseignement des mathématiques ou le vécu des mathophobes. *Revue des sciences de l'éducation, 15*(2), 193-218. doi: <http://doi.org/10.7202/900627ar>
- Gierl, M. J. (1997). *An investigation of the cognitive foundation underlying the rule-space model* (Thèse de doctorat non publiée). University of Illinois, Urbana-Champaign, IL. Extrait de <http://hdl.handle.net/2142/21041>
- 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, Urbana-Champaign, IL. Extrait de <http://search.proquest.com/dissertations/docview/305590285/1404F5D9FD55197C276/1?accountid=12543>
- Leighton, J. P., & Gierl, M. J. (2007). *Cognitive diagnostic assessment for education: Theory and applications*. Cambridge, MA: Cambridge University Press.
- Linacre, J. M. (1994). Sample size and item calibration stability. *Rasch Measurement Transactions, 7*(4), 328. Extrait de <http://www.rasch.org/rmt/rmt74m.htm>
- Lord, F. M. (1983). Small N justifies Rasch models. In D. J. Weiss (Ed), *New horizons in testing* (pp. 51-61). New York, NY: Academic Press.
- Loye, N. (2008). *Conditions d'élaboration de la matrice Q des modèles cognitifs et impact sur sa validité et sa fidélité* (Unpublished doctoral dissertation). University of Ottawa, Canada. Extrait de <http://search.proquest.com/dissertations/docview/304340176/1404F6140457A199D08/1?accountid=12543>
- Loye, N. (2010). 2010, odysée 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. In G. Raïche, K. Paquette-Côté, & D. Magis (Eds.), *Des mécanismes pour assurer la validité de l'interprétation de la mesure en éducation*, (volume 1, pp. 11-30). Ste-Foy, Québec : Presses de l'Université du Québec.

- Marcoux, G., Fagnant, A., Loye, N., & Ndinga, P. (sous presse). L'évaluation diagnostique des compétences scolaires. In C. Dierendonck, E. Loarer & B. Rey (Eds.), *L'évaluation des compétences en milieu scolaire et en milieu professionnel*. Bruxelles, Belgique: De Boeck.
- MELS (2003). *Politique d'évaluation des apprentissages*. Québec, QC: Gouvernement du Québec. Extrait de <http://www.mels.gouv.qc.ca/lancement/PEA/>
- MELS (2007). *Exploration de la formation professionnelle*. Québec, QC: Gouvernement du Québec. Extrait de [http://www.mels.gouv.qc.ca/sections/programmeformation/secondaire2/medias/07-00840\\_explorationFormProf.pdf](http://www.mels.gouv.qc.ca/sections/programmeformation/secondaire2/medias/07-00840_explorationFormProf.pdf)
- MELS (2008). *La formation professionnelle et technique au Québec-un aperçu*. Québec, QC: Gouvernement du Québec. Extrait de [http://www3.mels.gouv.qc.ca/fpt/ Bibliotheque/17-9828.pdf](http://www3.mels.gouv.qc.ca/fpt/Bibliotheque/17-9828.pdf)
- MELS (2010). *La formation professionnelle et technique au Québec-un aperçu*. Québec, QC: Gouvernement du Québec. Extrait de [http://www.mels.gouv.qc.ca/sections/publications/publications/FPT\\_FC/Formation\\_professionnelle\\_technique/LaFPTAuQuebec\\_UnApercu\\_2010\\_f.pdf](http://www.mels.gouv.qc.ca/sections/publications/publications/FPT_FC/Formation_professionnelle_technique/LaFPTAuQuebec_UnApercu_2010_f.pdf)
- Milewski, G. B., & Baron, P. A. (2002, April). *Extending DIF methods to inform aggregate reports on cognitive skills*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA. Extrait de <http://eric.ed.gov/?id=ED466712>
- Mundfrom, D. J., Shaw, D. G., & Ke, T. L. (2005). Minimum sample size recommendations for conducting factor analyses. *International Journal of Testing*, 5, 159-168. doi: [http://dx.doi.org/10.1207/s15327574ijt0502\\_4](http://dx.doi.org/10.1207/s15327574ijt0502_4)
- Pallant, J. F., & Tennant, A. (2007). An introduction to the Rasch model: An exemple using the Hospital anxiety and depression scale (HADS). *British journal of Clinical Psychology*, 46, 1-18. doi: <http://dx.doi.org/10.1348/014466506X96931>
- Ramp, M., Khan, F., Misajon, R. A., & Pallant, J. (2009). Rasch analysis of the multiple sclerosis impact scale (MSIS-29). *Health and Quality of Life Outcomes*, 7(58), 1-10. doi: <http://dx.doi.org/10.1186/1477-7525-7-58>
- Rupp, A. A., Templin, J., & Henson, R. J. (2010). *Diagnostic measurement: Theory, methods, and applications*. New York, NY: The Guilford Press.
- Shafer, L. S., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7(2), 147-177. doi: <http://dx.doi.org/10.1037%2F1082-989X.7.2.147>
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using Multivariate Statistics*. Boston, MA: Pearson Education.
- Tatsuoka, K. K. (2009). *Cognitive assessment: An introduction to the rule space method*. New York, NY: Routledge Taylor & Francis Group.
- Tobias, S. (1993). *Overcoming math anxiety, revised and expanded*. New York, NY: W.W. Norton & Company.
- Vanderwood, M., Ysseldyke, J., & Thurlow, M. (1993). *Consensus building: A process for selecting educational outcomes and indicators*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Extrait de <http://files.eric.ed.gov/fulltext/ED366171.pdf>

- Vukovic, R. K., Kieffer, M. J., Bailey, S. P., & Harari, R. R. (2013). Mathematic anxiety in young children: concurrent and longitudinal associations with mathematical performance. *Contemporary Educational Psychology*, 38(1), 1-10. doi: <http://dx.doi.org/10.1016/j.cedpsych.2012.09.001>
- Winter, J. C. F., Dodou, D., & Wieringa, P. A. (2009). Exploratory factor analysis with small sample sizes. *Multivariate Behavioral Research*, 44(2), 147-181. doi: <http://dx.doi.org/10.1080/00273170902794206>

Date de réception : 6 novembre 2012

Date de réception de la version finale : 4 septembre 2013

Date d'acceptation : 9 septembre 2013