



uOttawa

COMMENT RÉALISER DES ANALYSES DE RASCH AVEC DES ÉCHANTILLONS DE GRANDE TAILLE (N ≥ 1000) ?

Marwa Hamdan
Superviseur: Eric Dionne
Université d'Ottawa

INTRODUCTION

QU'EST-CE QUE LA MODÉLISATION RASCH

Ce type de modélisation permet essentiellement de situer les scores des sujets et des items sur une seule et même échelle de mesure. Le principe consiste à utiliser les scores bruts et de vérifier jusqu'à quel point ils se conforment aux données théoriques de la modélisation. Des statistiques d'ajustement permettent de déterminer jusqu'à quel niveau le modèle théorique et les scores empiriques s'ajustent bien.

BUT DE LA RECHERCHE

Depuis les années 1960, des analyses de Rasch sont réalisées pour modéliser les scores bruts à des tests ou instruments divers utilisés dans des domaines aussi variés que l'éducation, la psychologie ou encore l'éducation médicale. Des méthodologies ont été produites et éprouvées au cours des dernières décennies (Bond et Fox, 2007; Eckes, 2011; Engelhard, 2013; Smith et Smith, 2004) mais peu d'elles nous renseignent lorsque nous devons composer avec des échantillons de sujets de grande taille (N ≥ 1000). Le but de la recherche est de documenter les stratégies méthodologiques pour l'application de la modélisation Rasch avec des échantillons de grande taille ainsi que de mieux comprendre les difficultés répertoriées et les limites.

MÉTHODOLOGIE

La recherche a consisté en une revue de la littérature à l'aide d'outils de recherche bibliographiques ainsi que l'analyse de leurs contenus. Nous avons basé nos travaux sur les recommandations de Jesson, Matheson et Lacey (2011). Les articles les plus pertinents ont été identifiés à l'aide de certains mots-clés (ex: Rasch, dichotomous, polytomous, big data, large scale, infit, outfit) prédéterminés en utilisant les bases de données ERIC, MEDLINE, et PUBMED. Les informations ont été colligées dans un tableau Excel qui contenait la références sous format APA 6^{ème} édition, la base de donnée utilisée, les mots clés dans l'article, ainsi que du résumé de l'article et la méthodologie utilisée dans l'étude.

Un processus itératif a été mené pour la sélection des articles. Nous nous sommes également assurés d'obtenir un accord inter-juge adéquat quant au contenu répertorié. À partir des 39 articles trouvés, 7 ont été conservés pour l'étude dans cette recherche. Une comparaison de la méthodologie utilisée dans les articles choisis a permis de ressortir des catégories de modèles utilisés. Un tableau de résultats (ci-dessous) a été élaboré pour combiner les articles comparés selon le modélisation utilisée dans la méthode de recherche.

RÉSULTATS:

Modèle de Rasch	Article (référence)	Échantillon	Procédures spécifiques tenant compte de la taille de l'échantillon pour l'estimation des paramètres
Modèle de crédit partiel (Partial credit model)	Bonk, W.J.; Okey, G.J. (2003). A many-facet Rasch analysis of the second language group oral discussion task. <i>Language Testing</i> 89-110	2424 étudiants	Analyse à travers FACETS 3.17. L'administration de chaque test a été calibré et analysé de manière indépendante de l'autre. La notation a été entrée dans le le Modèle à échelle de notation (Rating Scale Model), et dans le Modèle de crédit partiel. Les deux modèles ont été comparés. Peu d'indice sur les problèmes d'ajustement des sujets.
	Yan, J.W. (1996). Examining Local Item Dependence Effects in a Large-Scale Science Assessment by a Rasch Partial Credit Model. Michigan State University, ProQuest, UMI Dissertations Publishing	10074 étudiants	Bigsteps a été employé pour réaliser els analyses. Aucune information sur les problèmes d'ajustement des sujets.
	Schulz, W.; Ainley, J.; Friedman, T.; Lietz, P. (2011) ICCS 2009 Latin American Report: Civic Knowledge and Attitudes among Lower-Secondary Students in Six Latin American Countries. International Association for the Evaluation of Educational Achievement	Jusqu'à 6576 étudiants	Les résultats d'estimation de vraisemblances ont été transformé à une échelle métrique avec une moyenne de 50 et un écart-type de 10. Aucune information sur les problèmes d'ajustement des sujets.
Modèle logistique à un paramètre (ML1P)	Coe, R. (2008) Comparability of GCSE Examinations in Different Subjects: An Application of the Rasch Model. <i>Oxford review of education</i> . 34.5: 609-636	610,258 étudiants	Deux méthodes ont été testées : (1) modéliser les scores avec tous les sujets et éliminer les scores des sujets qui ne s'ajustaient pas au modèle, (2) conserver les scores des sujets qui s'ajustaient bien au modèle et ajouter progressivement d'autres groupes de sujets. La deuxième méthode s'est avérée la plus intéressante.
	Kubinger, K.D.; Hohensinn, C.; Hofer, S.; Khorramdel, L.; Frebort, M.; Holoche-Ertl, S.; Reif, M.; Sonnleitner, P. (2011). Designing the test booklets for Rasch model calibration in a large-scale assessment with reference to numerous moderator variables and several ability dimension. <i>Educational Research and Evaluation</i> 483-495	1642 étudiants	Aucune information sur les problèmes d'ajustement des sujets. Les auteurs ne présentent aucune information pertinente sur la façon dont ils ont mené leurs analyses de Rasch.
Le modèle logistique multinomial à coefficients aléatoires	Tormakangas, K. (2011) Advantages of the Rasch Measurement Model in Analysing Educational Tests: An Applicator's Reflection. <i>Education research and evaluation</i> 15.5: 307-320.	3695 étudiants (Final study of CIVIS in 1999 and Pilot study of ICCS in 2008)	Aucune information sur les problèmes d'ajustement des sujets. Les auteurs ne présentent aucune information pertinente sur la façon dont ils ont mené leurs analyses de Rasch.
	Baghaei, P. (2012) The Application of Multidimensional Rasch Models in Large Scale Assessment and Validation: An Empirical Example. <i>Electronic Journal of Research in Educational Psychology</i> 10.1: 233-252	1021 étudiants	Programme ConQuest et estimation Expected A Posteriori (EAP)

INTERPRÉTATION

Parmi les articles choisis, il s'avère que la majorité des recherches entreprises avec les grands échantillons ont utilisé le Modèle de crédit partiel de Rasch. Deux recherches, ayant des échantillons de taille variant de 1500 à 3700 étudiants, ont utilisé le Modèle logistique à un paramètre (modèles dichotomiques de théorie de réponses aux items) et une recherche a employé un Modèle multidimensionnel de Rasch.

L'article de Coe (2008) a été le seul à décrire la méthodologie utilisée pour les problèmes d'ajustements des sujets.

BIBLIOGRAPHIE PARTIELLE

- Coe, R. (2008) Comparability of GCSE Examinations in Different Subjects: An Application of the Rasch Model. *Oxford review of education*. 34.5: 609-636
- Kubinger, K.D.; Hohensinn, C.; Hofer, S.; Khorramdel, L.; Frebort, M.; Holoche-Ertl, S.; Reif, M.; Sonnleitner, P. (2011). Designing the test booklets for Rasch model calibration in a large-scale assessment with reference to numerous moderator variables and several ability dimension. *Educational Research and Evaluation* 483-495
- Schulz, W.; Ainley, J.; Friedman, T.; Lietz, P. (2011) ICCS 2009 Latin American Report: Civic Knowledge and Attitudes among Lower-Secondary Students in Six Latin American Countries. International Association for the Evaluation of Educational Achievement
- Bonk, W.J.; Okey, G.J. (2003). A many-facet Rasch analysis of the second language group oral discussion task. *Language Testing* 89-110
- Yan, J.W. (1996). Examining Local Item Dependence Effects in a Large-Scale Science Assessment by a Rasch Partial Credit Model. Michigan State University, ProQuest, UMI Dissertations Publishing