

La psychologie différentielle au 21^{ème} siècle : nouvelles modélisations et applications



Sous la direction de
Julie Collange
Even Loarer
Todd Lubart

La psychologie différentielle au 21^{ème} siècle : nouvelles modélisations et applications

Sous la direction de

Julie Collange

Even Loarer

Todd Lubart



© Éditions Laboratoire Lati
ISBN : 978-2-9563911-0-4

Évaluation du fonctionnement différentiel des items et de l'influence de diverses variables sociodémographiques dans les subtests du WISC-IV

Sotta Kieng ([1,2](#)), Philippe Golay ([3,4](#)), & Thierry Lecerf (1,2)

Introduction

Pour l'adaptation française du WISC-IV – quatrième édition de l'Échelle d'Intelligence de Wechsler pour Enfants et Adolescents (Wechsler, 2005)–, une pré-expérimentation sur des échantillons de 220 enfants issus de régions françaises et 125 enfants belges ont abouti à la sélection des items possédant les qualités métriques requises.

Comme la validité de l'interprétation des scores est compromise par des items biaisés, les objectifs de cette étude sont d'évaluer l'influence des variables âge, sexe et professions des parents sur les items du WISC-IV pour un large échantillon d'enfants Suisse-Romands. Les proportions de variance expliquée par l'âge, le sexe et le statut socio-économique des parents seront examinées, ainsi que le fonctionnement différentiel des items.

Méthode

Sujets

L'échantillon se compose de 483 enfants (230 garçons) non consultants de 7 à 12 ans (âge moyen = 9.04 ans ; écart type = 1.31 an). Les enfants proviennent d'une vingtaine d'établissements scolaires du canton de Genève. Choisis au hasard, les sujets forment un échantillon relativement représentatif de la population d'enfants genevois au niveau de la variable sexe. En comparaison avec les statistiques du département de l'instruction publique (DIP) du Canton de Genève⁵, il y a 50% de garçons et 50% de filles dans les écoles publiques genevoises contre 48% de garçons et 52% de filles dans notre échantillon. Quant à la répartition des catégories socio-professionnelles des parents, l'échantillon comporte une surreprésentation d'enfants dont les parents appartiennent à la catégorie « Cadre supérieurs et dirigeants » (32% contre 20%) et une sous-représentation pour la catégorie « Ouvriers, divers et sans indication » (32% contre 35%). Nos critères d'exclusion étant la langue et le parcours scolaire, seuls les enfants francophones qui n'ont ni doublé ni sauté une ou plusieurs classes ont été retenus.

Instrument

Le WISC-IV a été standardisé sur un échantillon de 1'103 enfants âgés de 6 à 16 ans 11 mois. Cette batterie de tests se compose de 10 subtests obligatoires et de 5 subtests optionnels. En plus de fournir un QI Total (QIT), le WISC-IV propose également un Indice de Compréhension Verbale (ICV), un Indice de Raisonnement Perceptif (IRP), un Indice de Mémoire de Travail (IMT) et un Indice de Vitesse de Traitement (IVT).

Procédure

Dans le cadre de notre recherche, chaque enfant est vu en passation individuelle lors de 2 séances d'environ 40 minutes durant les heures scolaires. Les locaux mis à disposition se trouvent dans l'enceinte de l'école de l'enfant. Les dix subtests principaux du WISC-IV (Cubes, Similitudes, Mémoire des chiffres, Identification de concepts, Code, Vocabulaire, Séquence Lettres-Chiffres, Matrices, Compréhension et Symboles) sont administrés, ainsi que le subtest optionnel Complètement d'images.

Un questionnaire est envoyé aux parents et permet de recueillir les diverses informations sociodémographiques. Les professions sont classées en 10 catégories, allant de la catégorie 1 « Directeurs, cadre de direction et gérants » à la catégorie 10 « Sans emploi, au foyer ».

Les subtests de l'IMT et de l'IVT n'ont pas été analysés. D'une part, il est théoriquement peu justifié de supposer un fonctionnement

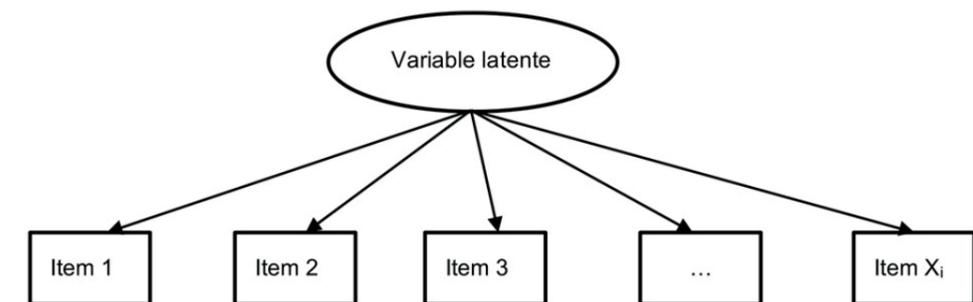
différentiel sur les items proposés par les subtests de l'IMT et d'autre part, il n'est pas possible d'ajuster un modèle de réponse à l'item sur les données des subtests de l'IVT. Il n'y a donc que 7 subtests dont la variable latente peut être estimée avec un modèle de mesure tiré des modèles de réponse à l'item. Les subtests avec des items à cotation dichotomique (Identification de concepts, Matrices et Complètement d'images) sont modélisés par un modèle de réponse à l'item logistique à 2 paramètres (2 PL). Quant aux subtests avec des items à cotation polytomique (Cubes, Similitudes, Vocabulaire et Compréhension), ils sont modélisés par un modèle de réponse à l'item gradué de Samejima (Samejima, 1970, 1997). Toutes les analyses ont été réalisées avec le programme Mplus 7.2 de Muthén & Muthén.

Résultats

Évaluation de l'unidimensionnalité des items

Avant de modéliser le fonctionnement des items, on procède à l'évaluation de l'unidimensionnalité des items de chaque subtest. Le modèle unidimensionnel postule que chaque subtest évalue une dimension psychologique principale. La dimension psychologique évaluée par le subtest prédit à elle seule les performances sur chacun des items (voir Figure 1). Selon ce postulat, la réussite (ou l'échec) aux différents items d'un subtest devrait uniquement dépendre du niveau du sujet sur la variable latente évaluée par le subtest, et donc ne pas être influencée par d'autres variables. Par exemple, pour le subtest Vocabulaire, on suppose que la dimension psychologique évaluée est principalement de la compréhension verbale ; le niveau du sujet sur le raisonnement ou la visualisation ne devrait pas influencer la réussite (l'échec) des items de ce subtest.

Figure 1 : Modèle unidimensionnel logistique à 2 paramètres (2 PL) ou gradué de Samejima.



Pour évaluer l'unidimensionnalité des items d'un subtest, on spécifie donc un modèle avec 1 variable latente réflexive qui explique les résultats sur chacune des variables manifestes comme illustré dans la Figure 1. En accord avec les recommandations de plusieurs auteurs (Hooper, Coughlan, & Mullen, 2008; Hu & Bentler, 1999; Yu, 2002)⁶, les indices d'ajustement (ou indices de fit) sont examinés afin d'estimer l'ajustement des données au modèle unidimensionnel. Les critères retenus sont : RMSEA < .06 et CFI > .95. Pour chaque subtest, les items de variances nulles ou faibles sont éliminés. Cette première série d'analyses montre un ajustement satisfaisant au modèle unidimensionnel pour Similitudes, Vocabulaire, Matrices et Compréhension. Étant donné l'ajustement insuffisant au modèle unidimensionnel, les épreuves Cubes, Identification de concepts et Complètement d'images sont retirées des analyses suivantes.

Différence de performances selon l'âge, le sexe et le statut socio-économique des parents

Le modèle unidimensionnel 2 PL de base est modifié avec l'introduction de la variable âge comme co-varié (voir Figure 2). Cette modification permet d'évaluer dans quelle proportion le score latent peut être expliqué par l'âge des participants. N'oublions pas que nous travaillons sur des scores bruts avec un échantillon d'enfants âgés de 7 à 12 ans, il est donc nécessaire d'estimer l'effet direct de l'âge.

¹ Université de Genève, Faculté de Psychologie et des Sciences de l'Éducation, 40 bd. du Pont d'Arve, 1205 Genève, Suisse.

² Filière Psychologie, Formation universitaire à distance, Suisse.

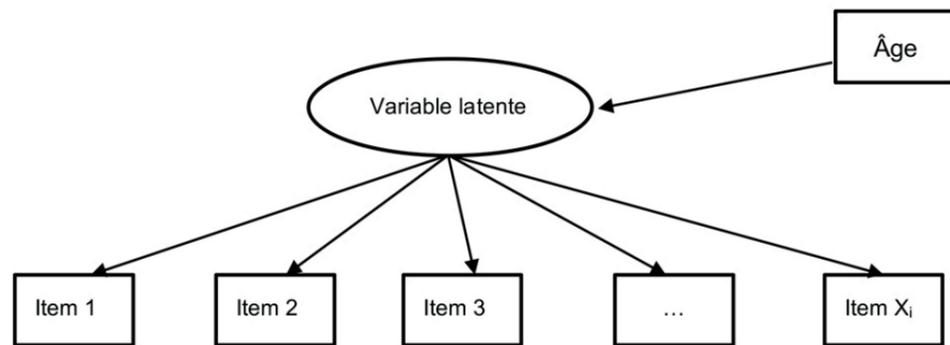
³ Université de Lausanne, Faculté des Sciences Sociales et Politiques, UNIL – Dorigny, Géopolis, CH-1015 Lausanne, Suisse.

⁴ Département de Psychiatrie, Centre Hospitalier Universitaire Vaudois (CHUV), Section de Psychiatrie sociale, 18 Place Chaudron, CH-1003 Lausanne.

⁵ Les statistiques sont celles de l'année scolaire 2014-2015

⁶ Comme sur de grands échantillons, un test de χ^2 non significatif ($p > .05$) est difficilement observable, nous n'avons pas tenu compte de ce critère.

Figure 2 : Modèle unidimensionnel avec la variable âge en co-varié.



Pour tous les subtests, les résultats montrent une influence directe de la variable âge sur le score latent (voir Tableau 1). Comme attendu, la variable âge est positivement liée à la réussite des items. En effet, un enfant plus âgé réussit un nombre plus élevé d'items qu'un enfant plus jeune, et donc obtiendra un score brut plus élevé. La variance expliquée par l'âge varie de 28.52% (MAT) à 45.56% (COM). On peut noter qu'elle est plus importante pour les subtests qui font appel au verbal.

Tableau 1 : Effet direct et variance expliquée par l'âge, le sexe et le statut socio-économique.

	Âge		Sexe		SES	
	Effet direct (β)	Variance expliquée	Effet direct (β)	Variance expliquée	Effet direct (β)	Variance expliquée
SIM	.631**	39.82%	-.023	-	-.284**	8.07%
VOC	.665**	44.22%	-.035	-	-.264**	6.97%
MAT	.534**	28.52%	.093*	0.86%	-.157**	2.46%
COM	.675**	45.56 %	-.077*	0.59 %	-.203**	4.12 %

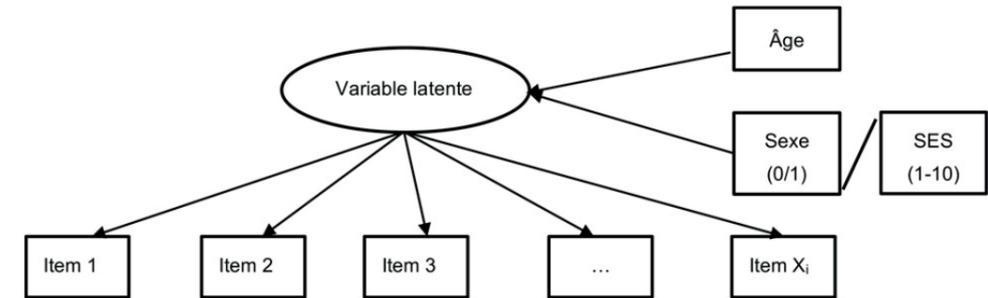
Note. SES = statut socio-économique ; SIM = Similitudes ; VOC = Vocabulaire ; MAT = Matrices ; COM = Compréhension.

* $p < .05$, ** $p < .01$.

Dans la suite des analyses, nous introduisons tour à tour comme co-varié supplémentaire les variables sexe de l'enfant, puis le statut socio-économique des parents. La variable âge est désormais contrôlée.

Nous testons l'effet direct de la variable sexe comme dans le modèle décrit dans la Figure 3. Les filles sont codées 1 et les garçons sont codés 0. Dans l'épreuve Compréhension, il y a une très faible influence de la variable sexe en faveur des garçons ($\beta = -.077$; voir Tableau 1). En revanche, une très faible influence de la variable sexe s'observe en faveur des filles dans l'épreuve des Matrices ($\beta = .093$).

Figure 3 : Modèle unidimensionnel avec les variables âge et sexe ou SES en co-varié.



La variable statut socio-économique des parents est construite à partir deux indicateurs : la profession du père et la profession de la mère. La situation professionnelle de chaque parent est assignée à l'une des 10 catégories. Plus la profession est socialement valorisée, plus la valeur du codage de la catégorie est basse. Le modèle testé est illustré par la Figure 3.

Les résultats indiquent un effet direct de la variable SES sur le score latent pour tous les subtests (voir Tableau 1). L'effet direct négatif de la variable SES suggère que les enfants ayant des parents avec une profession socialement valorisée ont des scores légèrement plus élevés que les enfants ayant des parents avec une profession socialement moins valorisée. La variance expliquée par la variable SES varie de 2.46% (MAT) à 8.07% (SIM).

Fonctionnement différentiel des items du WISC-IV

La dernière étape consiste à tester un éventuel fonctionnement différentiel des items. Pour ce faire, on examine les indices de modifications issus de la procédure d'estimation. Si un lien direct entre un co-varié et l'un des items d'un subtest est suggéré, cela signifie que l'item en question fonctionne différemment selon la modalité de l'une ou l'autre co-varié examiné (âge, sexe ou SES). En l'absence de biais, on s'attend à ce que la variable latente de la dimension évaluée par le subtest soit la seule et unique source de variations dans le score à l'item.

Les résultats vont dans le sens d'une absence de biais, puisqu'ils ne révèlent aucun fonctionnement différentiel des items pour les subtests analysés. En effet, une fois l'influence des variables âge, sexe ou statut socio-économique contrôlée, les capacités dans la dimension psychologique évaluée par un subtest prédisent à elles seules la réussite (ou l'échec) dans les items. Il n'y a donc pas d'effet indirect lié aux variables contrôlées.

Conclusion

Dans les bilans cognitifs, le psychologue s'appuie souvent sur l'interprétation de tests – tels que les populaires Échelles de Wechsler – pour poser des diagnostics et orienter les prises en charge. Étant donné les enjeux inhérents à de telles décisions, il est important que les différences observées sur les scores du test traduisent bel et bien des différences interindividuelles dans l'attribut psychologique, et non des différences liées à d'autres caractéristiques (ethnie, sexe ou milieu socio-économique par exemple). Nos résultats sur un échantillon de 483 enfants vont dans le sens d'une absence de biais sur les items des subtests Similitudes, Vocabulaire, Matrices et Compréhension.

Bibliographie

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, 6(1), 53-60.

Hu, L. t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. doi:10.1080/1070519909540118

Samejima, F. (1970). Estimation of latent ability using a response pattern of graded scores. *Psychometrika*, 35, 139-139. doi:10.1007/BF02290599

Samejima, F. (1997). Graded response model. In W. van der Linden & R. Hambleton (Eds.), *Handbook of Modern Item Response*

Theory (pp. 85-100): Springer New York.

Wechsler, D. (2005). *WISC-IV - Manuel d'interprétation*. Paris: Editions du Centre de Psychologie Appliquée.

Yu, C.-Y. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes*. University of California, Los Angeles, USA.

Exploring Different Assessment Techniques of the Test for Creative Thinking – Drawing Production

Christiane Kirsch ⁽¹⁾, Todd Lubart ⁽²⁾, & Claude Houssemand ⁽³⁾

Introduction

Presently, two different sets of creativity tasks are used, divergent-exploratory thinking tests and convergent-integrative thinking tests. The evaluation of creativity in the present research was based on a convergent-integrative measure, the Test for Creative Thinking-Drawing Production (TCT-DP). The TCT-DP consists in incorporating six dispersed geometric forms into one single drawing. Participants are actively invited to complete the drawing in an original way, within a restricted time frame. Drawing form B consists in exactly the same task as drawing form A, only that it is presented in a 180° rotation. The TCT-DP refers, according to Urban (2005), to a holistic gestalt-oriented conception of creativity and has been validated on a broad range of age, culture and ability groups. It proved itself as culture fair.

To evaluate the drawing, three distinctive scoring methods from different authors exist. The main aspect of the present research consists in comparing the three evaluation methods of the TCT-DP: (1) a traditional scoring system based on several sub scores (Urban, 1991, 2005; Urban & Jellen, 1996); (2) an originality score based on statistical frequency of ideas within the sample: statistical originality (SO) (Lubart, Pacteau, Jacquet, & Caroff, 2010); and (3) ratings given by trained judges according to the consensual assessment technique (CAT) (Amabile, 1982). Whereas these three scoring procedures were already set into comparison in former research done on school children (Lubart et al., 2010), the present study extends these results to the adult population in general and to specific sub-groups. It was analysed how the relationship between the scoring systems is modulated by the respective population: architects, social scientists and the general population. This will enable future research to fortify and specify the model of prediction of creativity in different populations, by always choosing the most appropriate evaluation technique of the TCT-DP.

According to Lubart et al. (2010), the ratings of both the traditional scoring system and the CAT include originality (novelty in the drawing content), reactivity/conformance to task constraints (use of the graphic elements provided) and complexity/quality of the drawing (mastery of artistic techniques). Through factor analyses, the presence of an Originality factor and an Adaptation factor in the traditional scoring system was empirically confirmed. This factor structure corresponds to the consensual creativity definition according to which creativity can be defined as the capacity to produce novel, original work that fits within contextual constraints (Lubart, Mouchiroud, Tordjman, & Zenasni, 2015).

Regardless of the obvious theoretical parallels existing between the traditional scoring system and the CAT, empirical findings concerning their association are slightly ambiguous. Whereas some studies revealed more significant correlations between both scoring techniques for drawing form A, others revealed more significant correlations for drawing form B.

With regard to drawing-form A, Dollinger, Urban and James (2004) revealed several significant correlations between the traditional scoring method and different scores of the CAT: the overall consensual assessment ratings ($r = .72, p < .001$), the overall details ratings, the overall gestalt ratings and the overall ratings made respectively by psychologist judges and artist judges. Whereas the details rating focused on artistic competencies and on elaboration quality of the single fragments, gestalt ratings were based on the level of integration of the isolated elements into one coherent picture. Hence, both scoring methods were found to meaningfully relate to each other for drawing-form A. On the other hand, Lubart et al. (2010) revealed significant to highly significant correlations between the traditional scoring method (OF/AF) and the CAT especially for drawing form B. For drawing-form A, only the correlation between the AF and the CAT was significant.

Despite the fact that findings concerning the relationship between the traditional method and the CAT are somewhat inconsistent, we hypothesize (1) significant correlations between the originality factor (OF)/ adaptation factor (AF) (traditional scoring method) and the CAT for drawing-form A and B. With regard to SO, Lubart et al. (2010) found highly significant correlations

¹ University of Luxembourg, ECCS Research Unit, Campus Belval, Maison des Sciences Humaines, 11, Porte des Sciences, L-4366 Esch-sur-Alzette, Luxembourg. E-mail: christijeanne@icloud.com

² Université Paris Descartes, Centre Henri Piéron, UFR Institut de Psychologie, 71 avenue Edouard Vaillant, 92774 Boulogne-Billancourt Cedex, France. Mel : todd.lubart@parisdescartes.fr

³ University of Luxembourg, ECCS Research Unit, Campus Belval, Maison des Sciences Humaines, 11, Porte des Sciences, L-4366 Esch-sur-Alzette, Luxembourg. E-mail: claud.houssemand@uni.lu

La psychologie différentielle au 21^{ème} siècle : nouvelles modélisations et applications



Sous la direction de
Julie Collange
Even Loarer
Todd Lubart

Contenu de l'ouvrage

La psychologie différentielle est entrée dans le 21^{ème} siècle. Quelles en sont les principales perspectives ? Cet ouvrage reprend une sélection de textes issus des 21^{ème} Journées Internationales de Psychologie Différentielle, qui témoignent de l'évolution des modélisations et des applications de la psychologie différentielle dans les domaines de la cognition, des émotions, de la personnalité, de l'éducation, de la formation et du travail. Les aspects différentiels pris en compte sont relatifs aux différences interindividuelles ainsi qu'à la variabilité intra individuelle et concernent tous les âges de la vie. Ils sont étudiés à l'aide de méthodes variées, telles que l'observation, l'expérimentation, la psychométrie, la simulation ou encore l'imagerie cérébrale.

On y trouve également une évocation de l'œuvre scientifique de Maurice Reuchlin (1920-2015), pionnier de la psychologie différentielle, soulignant la remarquable puissance de sa pensée et la grande modernité de ses travaux.

Cet ouvrage témoigne de la richesse et de la diversité des recherches en psychologie différentielle, de ses contributions au développement des connaissances fondamentales en psychologie mais également de son impact sur les pratiques dans de nombreux domaines de la vie quotidienne des individus.