

Sommaire

Sur la notion de complexité d'item : un modèle de réponse pour items complexes

Yvonnick Noël

Université Rennes 2, LP3C

MODEVAIIA, Saint Pierre de l'Isle, 9 juin 2023

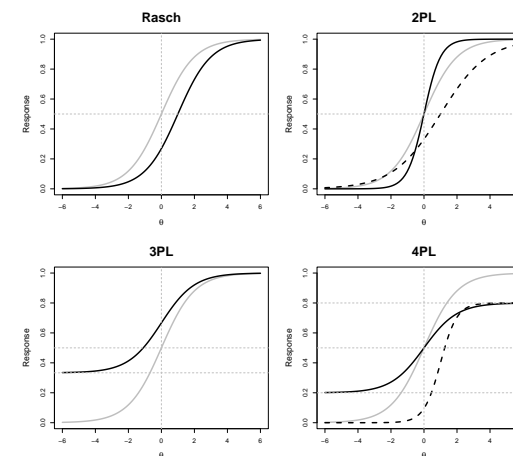


- 1 Les modèles de réponse logistiques
- 2 Notion de complexité d'item
 - Les travaux de Samejima
 - Un cadre en équations différentielles
- 3 Le modèle de Cascade de Logistiques Généralisées (COGL)
- 4 Applications
 - Compréhension de l'implicite des textes (TACIT)
 - Maîtrise de la langue arabe (CIMA)
 - Compétences en mathématiques (TIMSS)
- 5 Conclusions et perspectives

Les modèles de réponse logistiques

Les modèles symétriques classiques

Les modèles logistiques classiques



Propriété : les variations de ces fonctions sont de même amplitude à égale distance de leur point d'inflexion

La famille logistique des fonctions de réponse

- Le modèle de Rasch (Rasch, 1960) :

$$P(X_j = 1|\theta) = \frac{\exp(\theta - \delta_j)}{1 + \exp(\theta - \delta_j)} = 1 - \frac{1}{1 + \exp(\theta - \delta_j)}$$

- Le modèle logistique à deux paramètres (Birnbaum, 1968) :

$$P(X_j = 1|\theta) = \frac{\exp[\alpha_j(\theta - \delta_j)]}{1 + \exp[\alpha_j(\theta - \delta_j)]} = 1 - \frac{1}{1 + \exp[\alpha_j(\theta - \delta_j)]}$$

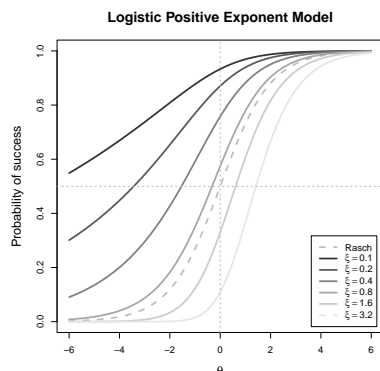
- Les modèles à 3 ($\lambda_j = 1$) et 4 paramètres (Birnbaum, 1968; Lord, 1980; Barton & Lord, 1980):

$$\begin{aligned} P(X_j = 1|\theta) &= \gamma_j + (\lambda_j - \gamma_j) \frac{\exp[\alpha_j(\theta - \delta_j)]}{1 + \exp[\alpha_j(\theta - \delta_j)]} \\ &= \lambda_j - \frac{\lambda_j - \gamma_j}{1 + \exp[\alpha_j(\theta - \delta_j)]} \end{aligned}$$

Notion de complexité d'item

- On parle d'**hétérogénéité** ou de **complexité** d'item quand la réussite à l'item est déterminée par un ensemble de sous-processus latents hétérogènes, articulés dans un système interactif (ex. la compréhension langagière).
- L'interaction ou l'intégration de ces sous-compétences, de manière dynamique et complexe, ne se laisse pas bien décomposer par des modèles multidimensionnels additifs.
- Certains auteurs (Samejima, 2000; Lee and Bolt, 2018) ont indiqué que la complexité d'item tendait à produire des **fonctions de réponse asymétriques**, rendant insuffisante la description en termes de difficulté et de discrimination.

Les travaux de Samejima (1969, 1972, 1999, 2000)

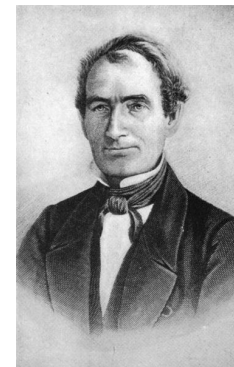


- Samejima (1972) a proposé le modèle candidat (*Logistic Positive Exponent Model*) :

$$P(X = 1|\theta) = \left[\frac{\exp[\alpha_j(\theta - \delta)]}{1 + \exp[\alpha_j(\theta - \delta)]} \right]^\xi$$

en donnant au paramètre exposant ξ l'interprétation de **paramètre d'accélération** (Samejima, 1999) ou de **complexité** (Samejima, 2000).

- Il augmente les **sources d'échec**.



Pierre-François Verhulst
(1804-1849)

Les travaux de Verhulst (1838)

- Le statisticien belge Verhulst (1838) cherchait une alternative bornée à l'exponentielle pour modéliser la **croissance des populations**.
- Il a proposé de la déduire comme solution de l'**équation différentielle de premier ordre** :

$$\frac{df(x)}{dx} = \alpha f(x)[1 - f(x)]$$

dont la solution est :

$$f(x) = \frac{\exp(\alpha x + C)}{1 + \exp(\alpha x + C)} = 1 - \frac{1}{1 + \exp(\alpha x + C)}$$

Une autre concept de cumul

- **Mécanisme** : on imagine que la réussite à un item Y est déterminée par la maîtrise d'un certain nombre de **sous-processus latents** (X_1, X_2, \dots) , causalement et séquentiellement **liés ensemble** $(X_1 \rightarrow X_2 \rightarrow \dots \rightarrow Y)$.
- Tous les processus et la réussite globale sur Y dépend d'une **unique compétence** θ .
- Cette idée de **séquence** de sous-processus est complémentaire des réflexions antérieures sur un impact en **parallèle** de sous-processus (Samejima, 2000; Lee & Bolt, 2018).

Un système dynamique aux dérivées partielles

- Considérons d'abord le cas simple où **la maîtrise d'un sous-processus** augmente avec la variable latente de compétence, tout en provoquant simultanément **une augmentation** dans le niveau de réussite observable à l'item Y , possiblement **avec un certain décalage de compétence**.
- On note x et y les fonctions d'intensité du sous-processus et de la réussite à l'item, respectivement. Une modélisation possible en **équations différentielles** de cette situation est :

$$\begin{cases} \frac{dx}{d\theta} = \alpha x(\theta)[1 - x(\theta)] \\ \frac{dy}{d\theta} = \beta x(\theta - \delta)[1 - y(\theta)] \end{cases}$$

avec α un **taux d'augmentation** pour la maîtrise du sous-processus x , β un **taux d'induction** du sous-processus x sur la réussite observable en y , et δ un paramètre de **décalage**.

Solutions

- Par **intégration**, la solution pour la fonction x est la logistique standard, et la solution pour y définit une classe de **fonctions logistiques généralisées** de la forme :

$$y(\theta) = 1 - \frac{D}{[1 + \exp[\alpha(\theta - \delta)]]^{\frac{\beta}{\alpha}}}$$

avec $D \in [0; 1]$.

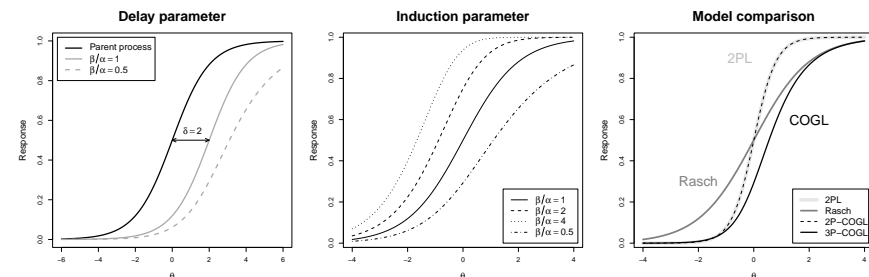
- Elle peut être généralisée encore en un modèle 4P en introduisant un **plafond** $\lambda \in [0; 1]$:

$$y(\theta) = \lambda - \frac{D}{[1 + \exp[\alpha(\theta - \delta)]]^{\frac{\beta}{\alpha}}}$$

intégrant ainsi les **modèles logistiques standard** (Rasch, 2P, 3P et 4P) comme cas particuliers (quand $\alpha = \beta$).

Propriétés

$$P(Y = 1|\theta) = 1 - \frac{1}{[1 + \exp[\alpha(\theta - \delta)]]^{\frac{\beta}{\alpha}}}$$



Relation au LPEM

- Avec le LPEM de Samejima (1972) :

$$P(X = 1|\theta) = \left[\frac{\exp[\alpha_j(\theta - \delta)]}{1 + \exp[\alpha_j(\theta - \delta)]} \right]^\xi = \left[\frac{1}{1 + \exp[-\alpha_j(\theta - \delta)]} \right]^\xi$$

le modèle COGL qui s'écrit :

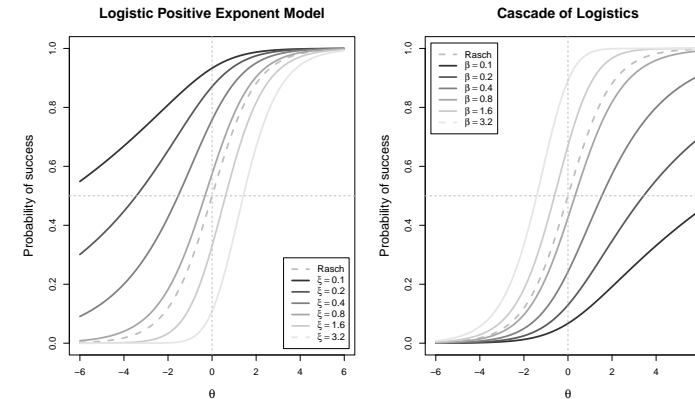
$$P(X = 1|\theta) = 1 - \frac{1}{[1 + \exp[\alpha(\theta - \delta)]]^{\frac{\beta}{\alpha}}} = 1 - \left[\frac{1}{1 + \exp[\alpha(\theta - \delta)]} \right]^{\frac{\beta}{\alpha}}$$

entretient la relation :

$$COGL(x) = 1 - LPEM(-x)$$

avec $\xi = \frac{\beta}{\alpha}$.

Relation au LPEM

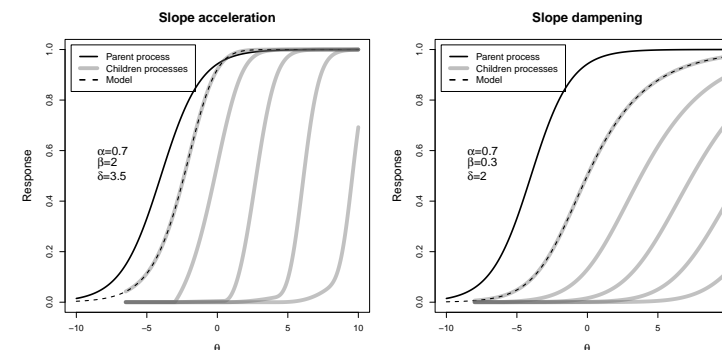


Interprétation

- L'exposant dans le LPEM **augmente la difficulté** dans les bas scores, et augmente la discrimination dans les hauts scores.
- L'exposant dans le COGL **diminue la difficulté** dans les hauts scores et augmente la discrimination dans les bas scores.
- Certains auteurs ont remarqué :
 - que le premier comportement (forte discrimination dans les items faciles) a tendance à apparaître dans les **épreuves linguistiques**,
 - que le second comportement (forte discrimination dans les items difficiles) a tendance à apparaître dans les **épreuves scientifiques et mathématiques**.

Cascade de logistiques généralisées (intégration numérique)

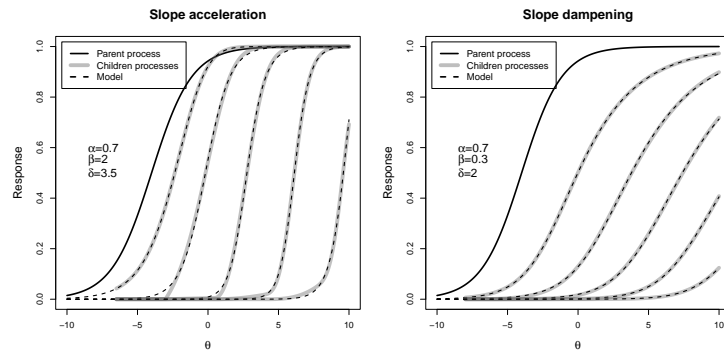
- En itérant (numériquement) ce processus d'induction, chaque **processus enfant devient un parent** à son tour :



- Mauvaise nouvelle : les intégrales sont **intraitables**...

Cascade de logistiques généralisées (approximation paramétrique)

- Mais si on utilise la solution à un seul processus, avec des **paramètres libres par item**, on obtient une approximation décente :



En résumé

- La fonction de réponse :

$$P(X = 1|\theta) = \lambda - \frac{D}{[1 + \exp[\alpha(\theta - \delta)]]^{\frac{\beta}{\alpha}}}$$

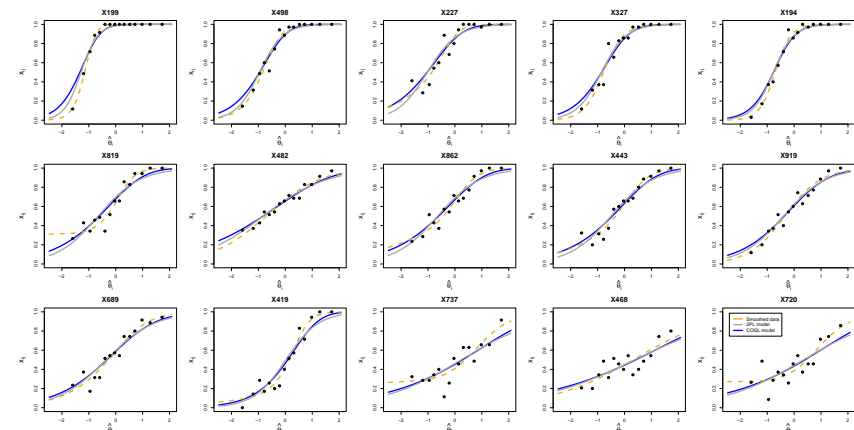
permet de modéliser avec une bonne approximation l'effet cumulé de **processus à croissance logistique placés en série**.

- L'usage de cette approximation nous oblige cependant à utiliser des paramètres totalement libres et indépendants par item, et nous **perdons l'interprétation** de ce qu'ils ont en commun.

Le projet TACIT

- TACIT (*Testing Adaptatif des Compétences Individuelles Transversales*) est une plate-forme en ligne d'évaluation et d'entraînement à la Compréhension Implicite de texte et de Vocabulaire, pour des élèves du CE1 à la 3ème : <http://www.tacit.fr> (Fanny De La Haye, Karine Lavandier, Olivier Le Bohec, Yvonnick Noel, Jeremie Nogues, Christophe Quaireau).
- Elle est basée sur le modèle de Rasch, et inclut un module de **Testing Adaptatif**.
- Parce que basée sur le Rasch, on ne s'attend pas à un bénéfice du COGL dans l'analyse.

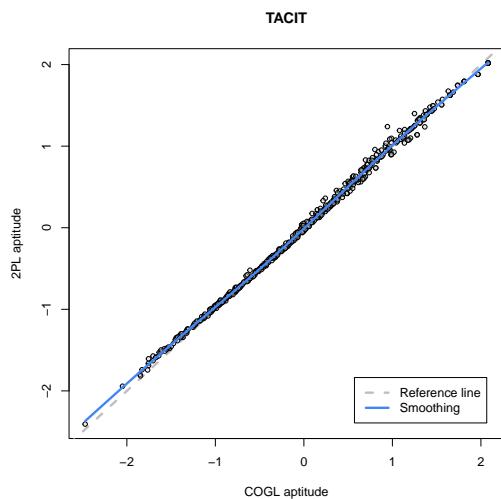
Application aux données TACIT



Légende : Données lissées, modèle 2PL , modèle COGL

Distorsions locales

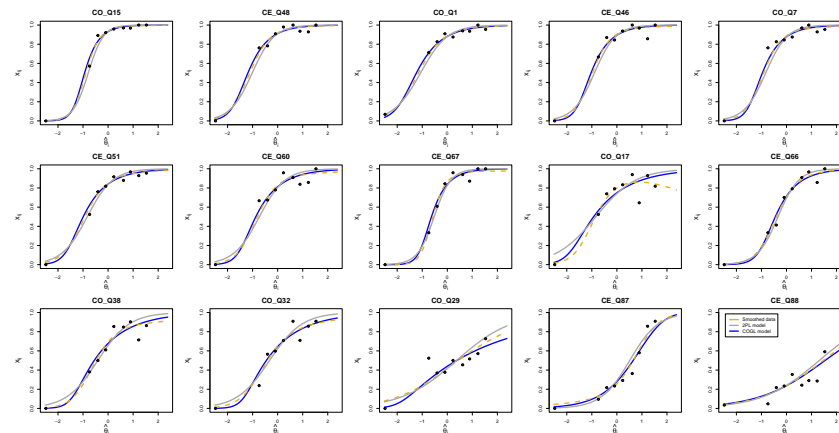
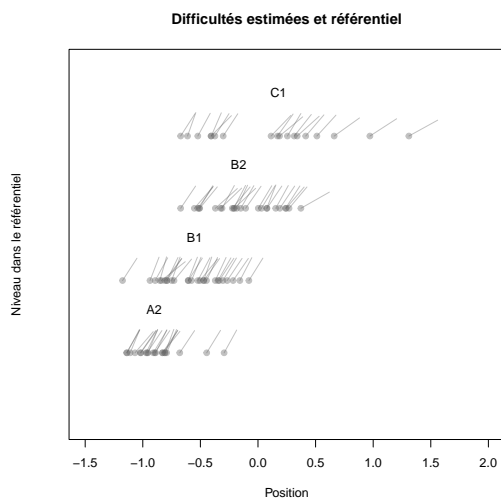
Développement d'un test de maîtrise de l'arabe



- Le **Certificat International de Maîtrise en Arabe (CIMA)** est une certification validant un niveau en arabe moderne standard.
- Il est développé par l'**Institut du Monde Arabe** à Paris sur le modèle du Test de Connaissance du Français (TCF) ou du TOEIC, avec l'expertise de France Éducation international.
- Il est passé chaque année par environ un millier de candidats, dans une douzaine de centres de langues dans le monde.
- **Objectif** : développer la partie psychométrique du test pour l'équipe du CIMA avec le **modèle de Rasch**.

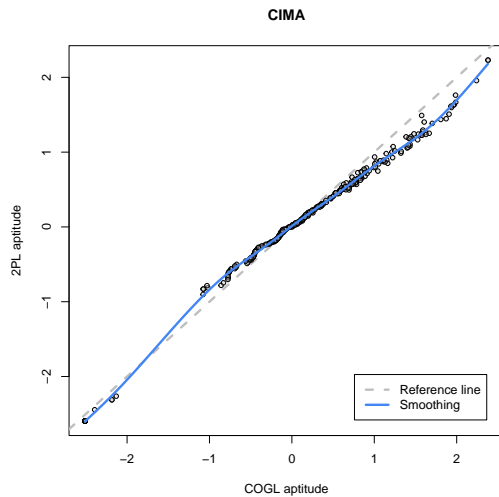
Test d'un modèle 2PL

Application aux données CIMA



Distorsions locales

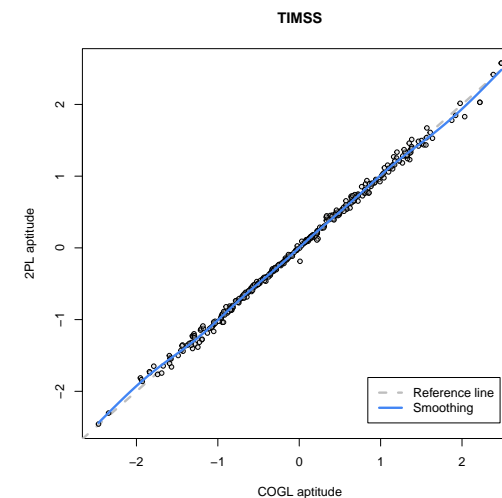
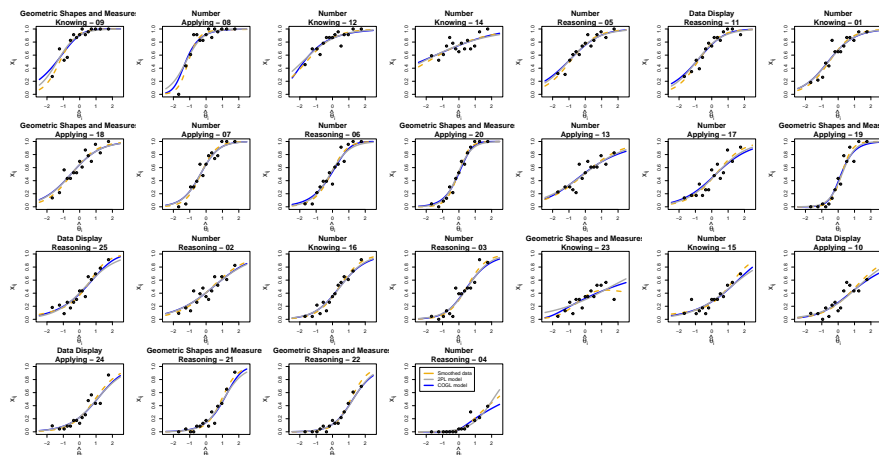
L'étude TIMSS



- TIMSS (*Trends in Mathematics and Science Study*) est une étude internationale comparative, qui mesure le niveau des connaissances des élèves de CM1 et de 4^{ème} **en mathématiques et en sciences**.
- Elle est organisée par une organisation internationale, indépendante, à but non lucratif qui conduit des études pédagogiques dans le monde entier : l'IEA (*International Association for the Evaluation of Educational Achievement*).
- Elle vise à interpréter les différences entre les systèmes éducatifs pour améliorer l'enseignement et l'apprentissage.

Application aux données TIMSS

Distorsions locales



- Le modèle COGL est à 3 paramètres et par conséquent **très flexible**. L'estimation des paramètres (séparés) est délicate (mais la prédiction excellente).
- En dépit de ces difficultés, le modèle COGL propose une nouvelle interprétation du phénomène de corrélation difficulté-discrimination, sous la forme d'un **cumul en cascade** de sous-processus en série.
- Une implémentation pour la famille des **Modèles de Réponse Beta** pour réponses continues bornées (Noël & Dauvier, 2007; Noël, 2014) est en développement.

Merci pour votre attention (yvonnick.noel@univ-rennes2.fr)