Mathématiques et sciences humaines
Mathematics and social sciences

184 | Hiver 2008
Varia

D'une échelle ordinale de Guttman à une échelle de rapports de Rasch

From a Guttman ordinal scale to a Rasch ratio scale

Daisy Bertrand, Abdessadek El Ahmadi et Christian Heuchenne

Édition électronique
URL : http://journals.openedition.org/msh/10959
DOI : 10.4000/msh.10959
ISSN : 1950-6821

Éditeur
Centre d'analyse et de mathématique sociales de l'EHESS

Édition imprimée
Date de publication : 31 décembre 2008
Pagination : 25-46
ISSN : 0987-6936

Référence électronique
Daisy Bertrand, Abdessadek El Ahmadi et Christian Heuchenne, « D'une échelle ordinale de Guttman à une échelle de rapports de Rasch », *Mathématiques et sciences humaines* [En ligne], 184 | Hiver 2008, mis en ligne le 25 février 2009, consulté le 19 avril 2019. URL : http://journals.openedition.org/msh/10959 ; DOI : 10.4000/msh.10959

Math. Sci. hum / Mathematics and Social Sciences (46e année, n° 184, *2008*(4), p. 25-46)

D'UNE ÉCHELLE ORDINALE DE GUTTMAN À UNE ÉCHELLE DE RAPPORTS DE RASCH

Daisy BERTRAND[1], Abdessadek El AHMADI[2], Christian HEUCHENNE†

RÉSUMÉ – *Cet article présente des modèles de mesure de plus en plus contraignants (échelle de Guttman dans sa version déterministe, version probabiliste de cette même échelle, échelle de rapports et enfin modèle de Rasch) avec en parallèle des conditions, nécessaires et suffisantes, de plus en plus astreignantes (Ferrers, forme probabiliste de Ferrers, …).*

MOTS-CLÉS – Biordre, Échelle ordinale, Échelle de rapports, Guttman, Mesurage, Rasch

SUMMARY – From a Guttman ordinal scale to a Rasch ratio scale
In this paper, increasingly restrictive measurement models are presented (Guttman scale, stochastic Guttman order, ratio scale and Rasch model) with, in parallel, their increasingly restrictive necessary and sufficient conditions (Ferrers relation, stochastic version of Ferrers, ...).

KEY-WORDS – Biorder, Guttman, Measurement, Ordinal scale, Ratio scale, Rasch

1. INTRODUCTION GÉNÉRALE : ÉCHELLE DE GUTTMAN

Le point de départ est l'observation de la réussite ou échec de sujets à des items. Techniquement, pour chaque couple (s, i) composé d'un sujet s et d'un item i est enregistré un et un seul événement :

s réussit i ou s ne réussit pas (rate) i

On dispose donc d'une relation empirique entre un ensemble S de sujets et un ensemble I d'items. On pourrait tout aussi bien étudier une relation entre S et I du type 'approuver' quand I est un ensemble d'opinions, 'exécuter' quand I est un ensemble de tâches, 'détecter' quand I est un ensemble de stimuli, etc. L'essentiel est que la réponse d'un sujet soit dichotomique ou dichotomisée.

---

[1] Laboratoire de psychologie cognitive, CNRS UMR 6146 - Université de Provence, 3 Place Victor Hugo, 13331 Marseille France, daisy.bertrand@univ-provence.fr

[2] Psychologie mathématique et statistique appliquée, UFR Psychologie & sciences de l'éducation – Université de Provence, 29 avenue Robert Schuman, 13621 Aix-en-Provence France, abdessadek.el-ahmadi@univ-provence.fr

† Le Professeur Christian Heuchenne (Université de Liège, Belgique) nous a brusquement quittés le 2 janvier 1998, laissant derrière lui une œuvre scientifique dont une partie est inédite ou inachevée. Que la publication de cet article, à partir de ses notes manuscrites, soit un hommage rendu à sa mémoire.

Pour simplifier l'exposé, les ensembles S et I seront supposés finis. L'extension à des ensembles infinis ne demande que quelques précautions supplémentaires qui n'ont d'intérêt que pour le mathématicien et seront omises par conséquent. Par exemple, avec 10 sujets et 6 items, la réussite est indiquée par 1 et l'échec par 0 dans la matrice

i_6	0	0	0	1	1	1	0	1	1	0
i_5	1	1	0	1	1	1	0	1	1	0
i_4	0	0	0	0	0	0	0	0	1	0
i_3	1	1	0	1	1	1	0	1	1	0
i_2	1	1	0	1	1	1	1	1	1	1
i_1	0	0	0	0	0	0	0	0	1	0
	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}

Tableau 1. Réponses observées pour dix sujets à six items dichotomiques

Communément, on a tendance à penser que s réussit i quand la compétence de s est assez forte pour surmonter la difficulté de i, au contraire que s échoue dans i lorsque la compétence de s est trop faible devant la difficulté de i. Formaliser cette idée banale, c'est supposer – à la suite de Guttman – que le sujet s a une certaine compétence, notée $\gamma(s)$, et l'item i une certaine difficulté, notée $\delta(i)$, de telle sorte que s réussisse i quand $\gamma(s) \geq \delta(i)$, mais rate i quand $\gamma(s) < \delta(i)$. Il y aurait donc, pour tout $s \in S$ et tout $i \in I$, les deux implications :

$\gamma(s) \geq \delta(i) \Rightarrow$ s réussit i et $\gamma(s) < \delta(i) \Rightarrow$ s ne réussit pas i

qui se résument en l'équivalence logique

s réussit i $\Leftrightarrow \gamma(s) \geq \delta(i)$ (M1)

Dès ce stade, il est important de remarquer que la relation 'réussir' entre S et I est observable tandis que γ et δ, applications respectivement de S et I vers l'ensemble R des nombres réels, sont des fonctions théoriques qu'il faut découvrir. Le sens concret de M1 est la possibilité de situer tous les sujets par compétence croissante et tous les items par difficulté croissante sur un même continuum orienté de telle manière qu'un sujet réussisse un item s'il est d'un côté de cet item et le rate s'il est de l'autre. De telles mesures conjointes constituent ce qu'on appelle, en sciences humaines, une échelle de Guttman ou un scalogramme [Guttman, 1944 ; Stouffer *et al.* 1950]. Le contenu mathématique de ce concept, qui a reçu depuis le nom de biordre, a été analysé dans les travaux de Ducamp et Falmagne [1969], de Doignon, Ducamp et Falmagne [1984] et de Falmagne [1985, 1988].

À quelle(s) condition(s) peut-on trouver des mesures γ et δ de la compétence et de la difficulté ayant la propriété (M1) ? La réponse, étonnamment simple, est la condition de Ferrers [voir Riguet, 1951 ; Monjardet, 1978] :

s'il existe un item réussi par le sujet s et raté par le sujet t,

alors tout item raté par s est aussi raté par t. (C11)

Un peu de réflexion montre que cette condition peut être énoncée sous la forme logiquement équivalente :

s'il existe un sujet qui réussit l'item j et rate l'item i,

alors tout sujet qui réussit i réussit j. (C12)

La condition de Ferrers revient à exiger que le Tableau 1 ne contienne aucune configuration de la forme $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ou $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Cette condition est nécessaire. En effet, s'il y a effectivement des fonctions γ de S vers R et δ de I vers R obéissant à (M1), et si i est un item réussi par s et raté par t, il vient $\gamma(s) \geq \delta(i)$ et $\delta(i) > \gamma(t)$, donc $\gamma(s) > \gamma(t)$. Si l'item j est raté par s, $\delta(j) > \gamma(s)$ entraîne $\delta(j) > \gamma(t)$, et j est aussi raté par t. Ce qui est remarquable et pas banal cette fois, c'est que la condition de Ferrers, d'apparences bien innocentes, est suffisante pour assurer l'existence de mesures γ et δ satisfaisant à (M1). La preuve générale se fonde sur une représentation numérique d'ensembles totalement ordonnés (théorème de Cantor, 1895) et n'est pas donnée ici [voir Birkhoff, 1967 ; Krantz *et al.* 1971]. Quand S et I sont finis, la démonstration, extrêmement instructive en elle-même, n'est qu'un long enchaînement de raisonnements.

Le noeud de l'affaire est que C11 ou C12 induisent en cachette une espèce d'ordre au sein des sujets et des items. Définissons une relation β dans la réunion des ensembles S et I, vérifiant la condition de Ferrers. Concrètement, s'agissant de sujets et d'items,

1. quand $s \in S$ et $i \in I$, $s\beta i$ signifie que le sujet s réussit l'item i.
2. quand $i \in I$ et $s \in S$, $i\beta s$ signifie que l'item i est raté par le sujet s.
3. quand $s \in S$ et $t \in S$, $s\beta t$ signifie que tout item raté par s l'est aussi par t.
4. quand $i \in I$ et $j \in I$, $i\beta j$ signifie que tout sujet qui réussit i réussit nécessairement j.

Cette définition posée, le résultat clé est que β est une relation réflexive, transitive et totale si la condition de Ferrers est satisfaite [voir Suppes et Zinnes, 1963 ; Ducamp et Falmagne, 1969]. La relation β rangeant ainsi conjointement sujets et items, l'interprétation générale de $a\beta b$ est que a 'domine' ou vaut b, $\forall\ a, b \in S \cup I$.

Dans l'exemple présenté, il est facile de vérifier l'une ou l'autre forme de la condition de Ferrers. L'ordonnancement par β conséquent se révèle en rangeant sujets et items par score croissant. Le score d'un sujet (resp. item) est le nombre d'items (resp. de sujets) qu'il a réussi (resp. qui l'ont réussi).

1	i_4	0	0	0	0	0	0	0	0	0	1
1	i_1	0	0	0	0	0	0	0	0	0	1
5	i_6	0	0	0	0	0	1	1	1	1	1
7	i_5	0	0	0	1	1	1	1	1	1	1
7	i_3	0	0	0	1	1	1	1	1	1	1
9	i_2	0	1	1	1	1	1	1	1	1	1
scores		s_3	s_7	s_{10}	s_1	s_2	s_4	s_5	s_6	s_8	s_9
	scores	0	1	1	3	3	4	4	4	4	6

Tableau 2. Matrice des réponses ordonnées de dix sujets à six items dichotomiques ; les 0 sont séparés des 1 par une 'frontière en escalier'.

La forme triangulaire de cette matrice (Tableau 2) est caractéristique d'une échelle de Guttman. Avec elle, les conditions (C11) ou (C12) sont immédiatement perçues.

Cette transparence n'apparaissait pas spontanément dans les données brutes où sujets et items étaient présentés dans un ordre arbitraire, sans référence à leurs compétences et difficultés encore inconnues.

Réexaminant la dernière matrice, on remarque ceci :

1. Certains sujets réagissent semblablement à tous les items et se valent donc de ce point de vue ; par exemple, s_7, s_{10} réussissent tous deux le seul item i_2 et manquent les autres ; s_4, s_5, s_6, s_8 réussissent tous i_2, i_3, i_5, i_6 et échouent sur i_1, i_4.
2. Certains items provoquent la même réponse chez tous les sujets et sont donc équivalents de ce point de vue ; ainsi i_3, i_5 sont tous deux ratés par s_3, s_7, s_{10} et réussis par tout autre sujet ; i_1, i_4 sont réussis par s_9 et manqués par les neuf autres répondants.

Ce phénomène général se marque par la décomposition de la relation β en deux parts, l'une symétrique qui est une équivalence, l'autre antisymétrique qui est un ordre strict [voir Suppes et Zinnes, 1963 ; Monjardet, 1978].

La part symétrique I_β de β se définit par :

a est lié par I_β à b quand simultanément $a\beta b$ et $b\beta a$, $\forall$ a, b $\in S \cup I$.

Notons déjà que ceci ne peut arriver quand a est un sujet et b un item, ou inversement, car $a\beta b$ (a réussit b) et $b\beta a$ (b est raté par a) sont contradictoires. La relation I_β ne peut donc exister qu'entre éléments de S ou entre éléments de I.

Quand a et b sont des sujets, $aI_\beta b$ signifie que tout item raté par a est raté par b (c'est $a\beta b$) et tout item raté par b est raté par a (c'est $b\beta a$). C'est dire formellement que a et b se comportent identiquement devant n'importe quel item.

Quand a et b sont des items, $aI_\beta b$ signifie que tout sujet qui réussit a réussit b (c'est $a\beta b$) et tout sujet qui réussit b réussit a (c'est $b\beta a$). C'est exprimer formellement que a et b suscitent la même réaction chez tous les sujets.

Il n'est donc pas étonnant que I_β soit une équivalence au sens mathématique du terme. En effet, il est évident que I_β est réflexive ($aI_\beta a$ pour tout a) et symétrique (si $aI_\beta b$, alors $bI_\beta a$), par définition. I_β est aussi transitive comme sa relation mère β : si $aI_\beta b$ et $bI_\beta c$, cela veut dire "$a\beta b$ et $b\beta a$" et "$b\beta c$ et $c\beta b$", ce qui revient à "$a\beta b$ et $b\beta c$" et "$c\beta b$ et $b\beta a$", lesquels entraînent par transitivité de β "$a\beta c$" et "$c\beta a$" qui n'est rien d'autre que $aI_\beta c$.

C'est un théorème classique qu'une équivalence partage son ensemble support en classes telles que deux éléments d'une même classe soient toujours liés alors que deux éléments de classes différentes ne soient jamais liés. Dans le contexte présent, I_β induit une partition de S en classes de sujets et une partition de I en classes d'items telles que :

- les sujets d'une même classe répondent à tous les items de la même manière ;
- de deux sujets de classes distinctes, l'un réussit et l'autre rate un item au moins ;
- les items d'une même classe provoquent une réponse identique chez tous les sujets ;
- deux items de classes distinctes entraînent l'un la réussite, l'autre l'échec, chez un sujet au moins.

Les classes apparaissent clairement dans l'exemple :

$\{s_3\}$ $\{s_7, s_{10}\}$ $\{s_1, s_2\}$ $\{s_4, s_5, s_6, s_8\}$ $\{s_9\}$ pour les sujets, $\{i_2\}$ $\{i_3, i_5\}$ $\{i_6\}$ $\{i_1, i_4\}$ pour les items.

L'autre composante de β est sa part antisymétrique D_β définie par :

$$\forall\ a, b \in S \cup I,$$

a est lié par D_β à b quand $a\beta b$ mais le lien réciproque $b\beta a$ ne tient pas.

Autrement dit, $aD_\beta b$ manifeste que a est lié à b par β sans lui être équivalent. Plus explicitement :

- quand $a \in S$ et $b \in I$, $aD_\beta b$ signifie que le sujet réussit l'item b (c'est $a\beta b$), ce qui entraîne que b n'est pas raté par a ($b\beta a$ est faux) ; dans l'exemple, $s_4D_\beta i_3$ parce que s_4 réussit i_3.
- quand $a \in I$ et $b \in S$, $aD_\beta b$ signifie que l'item a est raté par le sujet b (c'est $a\beta b$), ce qui entraîne que b ne réussit pas a ($b\beta a$ est faux) ; dans l'exemple, $i_6D_\beta s_2$ parce que i_6 est raté pas s_2.
- quand $a \in S$ et $b \in S$, $aD_\beta b$ signifie qu'il existe au moins un item raté par b et réussi par a ($b\beta a$ est faux), ce qui entraîne par (C11) que tout item raté par a est raté par b (c'est $a\beta b$) ; dans l'exemple, $s_1D_\beta s_3$ parce qu'il y a un item, tel i_5, raté par s_3 et réussi par s_1, et par conséquent tout item raté par s_1 l'est aussi par s_3.
- quand $a \in I$ et $b \in I$, $aD_\beta b$ signifie qu'il existe au moins un sujet qui réussit b et rate a ($b\beta a$ est faux), ce qui entraîne par (C12) que tout sujet qui réussit a réussit b (c'est $a\beta b$) ; dans l'exemple, $i_1D_\beta i_2$ parce qu'il y a un sujet, tel s_4, qui réussit i_2 et rate i_1, et donc tout sujet qui réussit i_1 réussit i_2.

Tandis que I_β rassemble entre eux sujets et items similaires, D_β oppose les éléments qui décidément se distinguent. Constatons encore que, puisque β est une relation totale ($a\beta b$ ou $b\beta a$), une et une seule des trois éventualités

$a\beta b$ et $b\beta a$ $\qquad$ $a\beta b$ et non($b\beta a$) $\qquad$ $b\beta a$ et non($a\beta b$)

peut se présenter, quel que soit le couple (a,b). En d'autres termes, on a la trichotomie suivante :

$aI_\beta b$ $\qquad$ $aD_\beta b$ $\qquad$ $bD_\beta a$

dont l'interprétation objective est

a est équivalent à b $\qquad$ a est supérieur à b $\qquad$ b est supérieur à a

En outre, D_β est un ordre strict au sens mathématique du mot. La preuve est courte : si l'on revient à sa définition, D_β est banalement antiréflexive ($aD_\beta a$ serait contradictoire) et antisymétrique ($aD_\beta b$ empêche $bD_\beta a$). La transitivité de D_β demande plus d'attention pour être établie : $aD_\beta b$ et $bD_\beta c$ impliquent en particulier $a\beta b$ et non($c\beta b$) qui, *a contrario* par la transitivité acquise de β, entraînent non($c\beta a$). Cette dernière proposition, avec $a\beta c$ inferée de $a\beta b$ et $b\beta c$, procure $aD_\beta c$.

La dernière étape à franchir est de voir que l'ordre D_β est en fait une relation totale entre les classes d'équivalence dégagées par I_β. Deux choses sont à établir :

1. si $aD_\beta b$, tout élément équivalent à a est lié par D_β à tout élément équivalent à b. En effet, soient a' équivalent à a et b' équivalent à b. $a'I_\beta a$ donne $a'\beta a$ et

$a\beta a'$, $b'I_\beta b$ donne $b'\beta b$ et $b\beta b'$, $aD_\beta b$ donne $a\beta b$ et non($b\beta a$). $a'\beta a$, $a\beta b$ et $b\beta b'$ entraînent $a'\beta b'$ par transitivité de β. Pour la même raison, $b\beta b'$, $b'\beta a'$ et $a'\beta a$ entraîneraient $b\beta a$; ainsi non($b\beta a$), $b\beta b'$ et $a'\beta a$ impliquent non($b'\beta a'$). Or $a'\beta b'$ et non($b'\beta a'$) signifient ensemble $a'D_\beta b'$.

2. si A et B sont des classes distinctes (de sujets ou d'items), tout élément de A est lié par D_β à tout élément de B ou réciproquement. En effet, soient $a \in A$ et $b \in B$ quelconques. Puisque les classes d'équivalence A et B sont distinctes, on ne peut avoir $aI_\beta b$; ne restent alors que les deux possibilités $aD_\beta b$ ou $bD_\beta a$.

Dans l'exemple, l'ordre strict et total D_β range les classes dans la liste suivante, les classes de sujets et d'items se succédant de manière intrinsèque :

$$\{s_3\} \quad \{i_2\} \quad \{s_7, s_{10}\} \quad \{i_3, i_5\} \quad \{s_1, s_2\} \quad \{i_6\} \quad \{s_4, s_5, s_6, s_8\} \quad \{i_1, i_4\} \quad \{s_9\}$$

Une classe A est liée par D_β à une classe B si et seulement si A suit B. En particulier, un sujet s réussit un item i si et seulement si s suit i.

Il est maintenant facile de trouver des mesures γ de la compétence des sujets et δ de la difficulté des items obéissant à (M1) : il suffit de disposer des nombres en suite croissante en regard des classes déjà ordonnées. Par exemple, la suite croissante quelconque (exemple 1 du Tableau 3) construit de bonnes mesures γ et δ.

En voici d'autres :

Ex.2 : à une classe de sujets peut être attribué leur score commun ;

Ex.3 : à une classe d'items peut être attribué l'opposé de leur score ;

Ex.4 : à une classe d'items peut être attribué l'inverse de leur score.

	$\{s_3\}$	$\{i_2\}$	$\{s_7, s_{10}\}$	$\{i_3, i_5\}$	$\{s_1, s_2\}$	$\{i_6\}$	$\{s_4, s_5, s_6, s_8\}$	$\{i_1, i_4\}$	$\{s_9\}$
Ex.1	1	1.5	2	3	4	4.5	5	6	7
Ex.2	0	0.5	1	2	3	3.5	4	5	6
Ex.3	-10	-9	-8	-7	-6	-5	-3	-1	1
Ex.4	1/10	1/9	1/8	1/7	1/6	1/5	1/3	1	0

Tableau 3. Différentes échelles ordinales rangeant conjointement sujets et items

Il s'agit d'échelles ordinales puisqu'une transformation strictement croissante des $\gamma(s)$ et $\delta(i)$ rend des mesures tout aussi admissibles. Par exemple, la transformation $f(x) = x-1$ fait passer de la première à la deuxième échelle.

De manière générale, si f est une fonction de R vers R strictement croissante ($x<y$ équivaut à $f(x) < f(y)$) et si γ et δ forment un modèle de Guttman, les mesures transformées $f(\gamma)$ et $f(\delta)$ en constituent un autre puisqu'elles satisfont encore à (M1) :

$$\text{s réussit i} \Leftrightarrow \gamma(s) \geq \delta(i) \Leftrightarrow f(\gamma(s)) \geq f(\delta(i)).$$

Synthétisons les étapes parcourues :

1. Si la relation 'réussir' de S vers I respecte la condition de Ferrers, elle engendre une relation β réflexive, transitive et totale dans la réunion de S et I.

2. β donne lieu à un ordre strict et total D_β entre classes d'équivalence dégagées par I_β.
3. Cet ordonnancement conjoint des sujets et items peut se représenter numériquement de diverses façons par une échelle ordinale de compétences et de difficultés.

On n'insistera jamais assez sur le point que la prémisse – la condition de Ferrers – n'est pas d'ordre mathématique mais ne peut que se constater dans la réalité par examen direct.

VERSION PROBABILISTE DE L'ÉCHELLE DE GUTTMAN

Si la modélisation de la réussite de sujets à des items veut être réaliste, la théorie précédente est trop abrupte et doit être assouplie. Il est exceptionnel qu'une échelle de Guttman, à cause de sa rigidité, s'applique parfaitement aux données expérimentales.

Il est naturel d'interpréter les écarts au modèle en concédant un caractère aléatoire à la relation empirique 'réussir' de S vers I. Ce caractère aléatoire est dû aux autres variables – non explicitement prises en compte comme la compétence des sujets et la difficulté des items – qui peuvent influencer la réussite ou l'échec ; on imagine aisément qu'elles sont complexes et nombreuses : humeur du sujet, environnement physique et social, mode de présentation de l'item, etc. [voir Linacre, 1992].

On conçoit donc que dans la situation où un sujet s est confronté à un item i, la réussite de i par s, au lieu d'être toujours réalisée quand $\gamma(s) \geq \delta(i)$, jamais quand $\gamma(s) < \delta(i)$, est gouvernée par une tendance floue : la réussite (et son contraire l'échec) a une certaine probabilité de survenir. Désormais ce n'est plus le fait de réussir, mais les chances de réussir qui seront fonctions de $\gamma(s)$ et $\delta(i)$. La probabilité que s réussisse i, notée $\pi(s,i)$, dépend de la compétence de s comme de la difficulté de i.

Pour illustrer ce propos, supposons que dans l'exemple précédent les probabilités $\pi(s,i)$ soient présentées dans la matrice suivante :

i_4	0	0	0.2	0.2	0.4	0.4	0.5	0.5	0.6	0.6
i_1	0	0	0.2	0.3	0.4	0.5	0.5	0.6	0.6	0.7
i_6	0	0	0.2	0.3	0.5	0.5	0.5	0.6	0.7	0.7
i_5	0	0	0.2	0.3	0.5	0.5	0.6	0.7	0.8	0.8
i_3	0	0.1	0.3	0.4	0.6	0.6	0.7	0.8	0.9	0.9
i_2	0	0.1	0.3	0.4	0.6	0.7	0.7	0.8	1	1
	s_3	s_7	s_{10}	s_1	s_2	s_4	s_5	s_6	s_8	s_9

Tableau 4. Version probabiliste de la matrice des réponses ordonnées
de dix sujets à six items

Si la compétence de s est faible devant la difficulté de i, $\pi(s,i)$ est petit, tel $\pi(s_1,i_1)=0.3$, et il y a plus de chances d'observer un échec qu'une réussite. À l'inverse, si s est plus compétent que i est difficile, $\pi(s,i)$ est grand, tel $\pi(s_8,i_5)=0.8$, et s réussit plus souvent i qu'il le rate. À la limite, quand $\pi(s,i)$ est un, tel $\pi(s_9,i_2)$, il est certain que s réussit i. Mais si $0<\pi(s,i)<1$, les deux éventualités restent possibles, même si l'une est plus plausible que l'autre : il y aura donc une grande variation dans les données qu'un même modèle peut engendrer.

Dans l'exemple, les observations consisteront en une matrice du type :

i_4	0	0	?	?	?	?	?	?	?	?
i_1	0	0	?	?	?	?	?	?	?	?
i_6	0	0	?	?	?	?	?	?	?	?
i_5	0	0	?	?	?	?	?	?	?	?
i_3	0	?	?	?	?	?	?	?	?	?
i_2	0	?	?	?	?	?	?	?	1	1
	s_3	s_7	s_{10}	s_1	s_2	s_4	s_5	s_6	s_8	s_9

Tableau 5. Matrice des réalisations possibles

La zone d'incertitude formée par les ? permet réussite 1 ou échec 0. L'échelle de Guttman étudiée ci avant est une des nombreuses réalisations possibles.

Mais surtout le nouveau modèle autorise des configurations qui ne sont pas des échelles de Guttman, telle la matrice :

1	i_4	0	0	0	0	0	0	0	0	0	1
2	i_1	0	0	0	0	0	0	1	0	1	0
4	i_6	0	0	0	0	0	1	0	1	1	1
7	i_5	0	0	0	1	1	1	1	1	1	1
8	i_3	0	0	1	1	1	1	1	1	1	1
8	i_2	0	1	1	0	1	1	1	1	1	1
scores		s_3	s_7	s_{10}	s_1	s_2	s_4	s_5	s_6	s_8	s_9
	scores	0	1	2	2	3	4	4	4	5	5

Tableau 6. Une des réalisations possibles, contredisant l'échelle de Guttman

Bien que d'allure presque triangulaire, il y a quelques écarts au modèle de Guttman. Rappelons que la disposition $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ou $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ contrevient à (C11) ou (C12), condition nécessaire et suffisante d'un scalogramme. Cette matrice ne constitue cependant pas un événement exceptionnellement rare ; par exemple, la réussite de s_7 à i_2 et l'échec de s_9 à i_1, bien que de petites probabilités $\pi(s_5,i_2)=0.1$ et $1-\pi(s_9,i_1)=0.3$, peuvent néanmoins advenir.

On est donc conduit à poser que $\pi(s,i)$ est une certaine fonction de la compétence de s et de la difficulté de i :

$$\pi(s,i) = g\{\gamma(s), \delta(i)\} \tag{M2}$$

La fonction g ne peut être entièrement quelconque. Si la compétence des sujets augmente, la probabilité de réussir tel item augmente : pour tout $i \in I$, $\gamma(s) \leq \gamma(t)$ implique $\pi(s,i) = g\{\gamma(s), \delta(i)\} \leq g\{\gamma(t), \delta(i)\} = \pi(t,i)$.

De même, si la difficulté des items augmente, la probabilité de réussite diminue : pour tout $s \in S$, $\delta(i) \leq \delta(j)$ entraîne $\pi(s,i) = g\{\gamma(s), \delta(i)\} \geq g\{\gamma(s), \delta(j)\} = \pi(s,j)$.

Donc la fonction $(x,y) \rightarrow g(x,y)$ doit être croissante en x et décroissante en y. Le problème est maintenant de trouver une telle fonction g de R^2 vers [0, 1] accompagnée d'applications γ de S vers R et δ de I vers R telles que $\pi(s,i) = g\{\gamma(s), \delta(i)\}$.

La question s'est considérablement déplacée par rapport à (M1) : d'une version déterministe de l'échelle de Guttman, on passe à une version assouplie, techniquement probabiliste de l'échelle de Guttman [voir Andrich, 1985 ; Falmagne, 1985, p. 38 ; Linacre, 1992].

La double condition suivante est alors nécessaire et suffisante pour fournir (M2) :

S'il existe un item i avec $\pi(t,i) < \pi(s,i)$, $\pi(t,j) \leq \pi(s,j)$ pour tout item j (C21)

S'il existe un sujet s avec $\pi(s,i) < \pi(s,j)$, $\pi(t,i) \leq \pi(t,j)$ pour tout sujet t (C22)

La nécessité de (C21) est vite aperçue. Puisque g est croissante en son premier argument, s'il y a $\delta(i)$ avec $\pi(t,i) = g\{\gamma(t),\delta(i)\} < g\{\gamma(s),\delta(i)\} = \pi(s,i)$, cela exige $\gamma(t) < \gamma(s)$, donc $g\{\gamma(t),\delta(j)\} \leq g\{\gamma(s),\delta(j)\}$ quel que soit $\delta(j)$. Tenant compte de la représentation (M2), c'est (C21).

Pour obtenir (C22) à partir de (M2), la décroissance de g en son second argument est semblablement exploitée.

L'analogie de forme de (C21) et (C22) avec les versions (C11) et (C12) de la condition de Ferrers mérite d'être notée. L'unidimensionnalité des sujets et des items est encore mise en évidence par (C21), mais cette fois sous forme probabiliste. Si le sujet t est moins fort en probabilité que le sujet s dans un item i, t est moins fort en probabilité que s sur tous les items. C'est affirmer autrement que les divers items ne sont que des manifestations d'un même comportement et qu'ils jouent de manière égale vis-à-vis de n'importe quel sujet. (C22) s'interprète de manière analogue.

Les conditions (C21) et (C22) se vérifient directement dans la matrice reprise dans le Tableau 4. S'il se trouve une ligne i où $\pi(t,i)<\pi(s,i)$, dans tout autre ligne j, $\pi(t,j)\leq\pi(s,j)$ est constaté ; s'il y a une colonne s où $\pi(s,i)<\pi(s,j)$, dans tout autre colonne t tient $\pi(t,i)\leq\pi(t,j)$. Dans cet exemple, la validité du modèle (M2) est donc certifiée. γ et δ décrits en marges de la matrice suivante (Tableau 7) illustrent ce fait ; il reste à prendre acte que la fonction g définie par

$$x \text{ en colonne, } y \text{ en ligne } \rightarrow g(x,y) \text{ de la matrice}$$

est bien croissante en x et décroissante en y.

Plus fondamentalement, le modèle (M2) correspond exactement à un agrégat d'échelles de Guttman à divers seuils de probabilité $\pi^* \in]0, 1[$.

5	i_4	0.0	0.0	0.2	0.2	0.4	0.4	0.5	0.5	0.6	0.6
3	i_1	0.0	0.0	0.2	0.3	0.4	0.5	0.5	0.6	0.6	0.7
1	i_6	0.0	0.0	0.2	0.3	0.5	0.5	0.5	0.6	0.7	0.7
0	i_5	0.0	0.0	0.2	0.3	0.5	0.5	0.6	0.7	0.8	0.8
-1	i_3	0.0	0.1	0.3	0.4	0.6	0.6	0.7	0.8	0.9	0.9
-2	i_2	0.0	0.1	0.3	0.4	0.6	0.7	0.7	0.8	1.0	1.0
δ(i)		s_3	s_7	s_{10}	s_1	s_2	s_4	s_5	s_6	s_8	s_9
	γ(s)	-6	-5	-4	-3	2	4	6	7	8	9

Tableau 7. Version probabiliste d'une échelle de Guttman

Créons une relation fictive « réussir à π* » en remplaçant π(s,i) par 1 quand π(s,i) ≥ π*, par 0 quand π(s,i) < π*. Cette relation de S vers I, obtenue par dichotomisation des π(s,i), doit être entendue comme « avoir une probabilité de réussite au moins égale à π* ». Dès lors, (C21) entraîne que la relation ainsi définie obéit à la version (C11) de la condition de Ferrers (et (C22) implique semblablement la version (C12) pour « réussir à π* »).

En effet, s'il existe i « réussi à π* » par s mais « raté à π* » par t, cela veut π(s,i) ≥ π* et π(t,i) < π*. Puisque π(t,i) < π(s,i), de (C21) découle π(t,j) ≤ π(s,j) pour tout j ∈ I. Ainsi, quand j est « raté à π* » par s, c'est-à-dire π(s,j) < π*, *a fortiori* π(t,j) < π*, soit j est « raté à π* » par t. On a prouvé que « réussir à π* » vérifie la condition (C11).

Au total, la relation « réussir à π* » engendre une échelle de Guttman pour π* arbitrairement choisi entre 0 et 1. C'est ce que l'on observe dans l'exemple suivant où des configurations triangulaires sont exhibées par diverses valeurs de π* :

			π*= 0.05		π*= 0.25		π*= 0.50					
5	i_4	0	0	.2	.2	.4	.4	.5	.5	.6	.6	π*=0.65
3	i_1	0	0	.2	.3	.4	.5	.5	.6	.6	.7	
1	i_6	0	0	.2	.3	.5	.5	.5	.6	.7	.7	
0	i_5	0	0	.2	.3	.5	.5	.6	.7	.8	.8	
-1	i_3	0	.1	.3	.4	.6	.6	.7	.8	.9	.9	π*=0.95
-2	i_2	0	.1	.3	.4	.6	.7	.7	.8	1	1	
δ(i)		s_3	s_7	s_{10}	s_1	s_2	s_4	s_5	s_6	s_8	s_9	
	γ(s)	-6	-5	-4	-3	2	4	6	7	8	9	

Tableau 8. Échelles de Guttman engendrées pour diverses valeurs critiques de π*

La fonction g étant croissante en γ(s) et décroissante en δ(i), il est clair que γ et δ ne sont encore que des échelles ordinales. Profitant de cet arbitraire, on peut déplacer l'une d'elles par une transformation strictement croissante en sorte que g(γ(s),δ(i)) ≥ 0.5 quand γ(s) ≥ δ(i) et g(γ(s), δ(i)) < 0.5 quand γ(s) < δ(i). C'est ce qui a été réalisé d'office dans l'exemple. L'inégalité γ(s) ≥ δ(i) devient alors équivalente à π(s,i) ≥ 0.50, ce qui montre lumineusement encore que le modèle présent (M2) est un avatar probabiliste d'une échelle de Guttman. Si la compétence de s domine la difficulté de i (dans l'exemple, $\gamma(s_6)=7 > 3=\delta(i_1)$, s réussit i à π* = 0.50 (dans l'exemple, $\pi(s_6,i_1) = 0.60$) ; inversement, si la difficulté de i surpasse la compétence de s (dans l'exemple, $\delta(i_2) = -2 > -3 = \gamma(s_1)$), s rate i à π* = 0.50 (dans l'exemple, $\pi(s_1,i_2) = 0.40$).

En résumé, le modèle M2 n'est qu'une version probabiliste, stochastique, de l'échelle de Guttman car, d'une part, C21 et C22 sont des formes probabilistes de C11 et C12, d'autre part, si on y fige un seuil π^* et si on définit conventionnellement la réussite de s sur i par $\pi(s,i) \geq \pi^*$, un modèle M1 est obtenu quel que soit la probabilité critique π^* entre 0 et 1.

ÉCHELLE DE RAPPORTS

Un modèle un peu plus exigeant précise que la probabilité de réussite ne dépend que du rapport entre compétence et difficulté mesurées positivement. Il s'exprime par

$$\pi(s,i) = F\{\gamma(s)/\delta(i)\} \qquad \text{(M3)}$$

où F est une fonction croissante de R+, ensemble des réels positifs, vers [0, 1], γ une application de S vers R+ et δ une application de I vers R+.

Ce modèle est un cas particulier du précédent puisqu'en posant $g(x,y) = F(x/y)$ en (M2), g sera croissante en x et décroissante en y. Il n'est donc pas étonnant qu'une condition nécessaire et suffisante de la représentation des probabilités (M3) est l'adjonction à (C21) et (C22) d'une condition supplémentaire.

Celle-ci s'énonce : pour tous $s, t, u \in S$ et tous $i, j, k \in I$

$$\pi(s,j) < \pi(t,k) \text{ et } \pi(t,i) < \pi(u,j) \text{ entraînent } \pi(s,i) \leq \pi(u,k) \qquad \text{(C3)}$$

Il est facile de voir que (C3) est nécessitée par (M3).

Si $\pi(s,j) = F\{\gamma(s)/\delta(j)\} < F\{\gamma(t)/\delta(k)\} = \pi(t,k)$ et $\pi(t,i) = F\{\gamma(t)/\delta(i)\} < F\{\gamma(u)/\delta(j)\} = \pi(u,j)$, la croissance de F demande que

$$\gamma(s)/\delta(j) < \gamma(t)/\delta(k) \text{ et } \gamma(t)/\delta(i) < \gamma(u)/\delta(j)$$

Par multiplication des deux dernières inégalités, on trouve

$$(\gamma(s)\gamma(t))/(\delta(j)\delta(i)) < (\gamma(t)\gamma(u))/(\delta(k)\delta(j))$$

d'où l'on tire

$$\gamma(s)/\delta(i) < \gamma(u)/\delta(k)$$

qui fournit enfin

$$\pi(s,i) = F\{\gamma(s)/\delta(i)\} \leq F\{\gamma(u)/\delta(k)\} = \pi(u,k).$$

Une fois encore, ce qui est intéressant, c'est que (C21), (C22), (C3) suffisent à assurer l'existence de fonctions F, γ, δ obéissant à (M3).

Dans la matrice (Tableau 4) des $\pi(s,i)$, (C21) et (C22) ont déjà été vérifiées et (C3) peut l'être maintenant avec beaucoup de patience. Des mesures γ, δ convenables sont par exemple données dans les marges de la matrice suivante (Tableau 9) :

14	i_4	0.00	0.14	0.42	0.57	0.85	0.92	1.00	1.14	1.28	1.35
13	i_1	0.00	0.15	0.46	0.61	0.92	1.00	1.07	1.23	1.38	1.46
12	i_6	0.00	0.16	0.50	0.66	1.00	1.08	1.16	1.33	1.50	1.58
11	i_5	0.00	0.18	0.54	0.72	1.09	1.18	1.27	1.45	1.63	1.72
10	i_3	0.00	0.20	0.60	0.80	1.20	1.30	1.40	1.60	1.80	1.90
9	i_2	0.00	0.22	0.66	0.88	1.33	1.44	1.55	1.77	2.00	2.11
δ(i)		s_3	s_7	s_{10}	s_1	s_2	s_4	s_5	s_6	s_8	s_9
	γ(s)	0	2	6	8	12	13	14	16	18	19

Tableau 9. Matrice contenant les rapports γ(s)/δ(i).

La matrice elle-même contient les γ(s)/δ(i). Il est facile de voir qu'en divisant ses composantes par 2, puis arrondissant au dixième inférieur, les π(s,i) de la matrice reprise dans le Tableau 4 sont exactement obtenus ; ainsi 14/11 = 1.273 de (s_5,i_5) donne 0.636, puis 0.6 = $\pi(s_5,i_5)$. Or la fonction F : z ∈ [0, 2.2] → "arrondi de z/2 au dixième inférieur" est croissante de R+ vers [0, 1]. On dispose bien d'une représentation des π(s,i) par F{γ(s)/δ(i)}.

La seule exigence sur F étant sa croissance, on peut toujours la construire de telle sorte que z ≥ 1 équivaille à F(z) ≥ 0.50. Dans ce cas, γ(s) ≥ δ(i) équivaut à π(s,i) = F{γ(s)/δ(i)} ≥ 0.50, comme c'était possible dans le modèle précédent. C'est ce qui a été fait ici aussi, comme on peut le contrôler (voir le résultat dans le Tableau 10).

14	i_4	0.0	0.0	0.2	0.2	0.4	0.4	0.5	0.5	0.6	0.6
13	i_1	0.0	0.0	0.2	0.3	0.4	0.5	0.5	0.6	0.6	0.7
12	i_6	0.0	0.0	0.2	0.3	0.5	0.5	0.5	0.6	0.7	0.7
11	i_5	0.0	0.0	0.2	0.3	0.5	0.5	0.6	0.7	0.8	0.8
10	i_3	0.0	0.1	0.3	0.4	0.6	0.6	0.7	0.8	0.9	0.9
9	i_2	0.0	0.1	0.3	0.4	0.6	0.7	0.7	0.8	1.0	1.0
δ(i)		s_3	s_7	s_{10}	s_1	s_2	s_4	s_5	s_6	s_8	s_9
	γ(s)	0	2	6	8	12	13	14	16	18	19

Tableau 10. Matrice correspondant à une échelle de rapports

En positionnant les items dans la liste des sujets, on peut conclure que la compétence de s est inférieure à la difficulté de i (γ(s) < δ(i)) si et seulement si s a moins de chances de réussir i que de le rater (π(s,i) < 0.50).

Quand F a été spécifiée, la seule liberté qui reste aux γ(s) et δ(i) est d'être multipliés par une constante ; par exemple, multiplier les compétences et difficultés précédentes par 11.37 ou les diviser par 5 ne change rien à l'affaire car les rapports γ(s)/δ(i) sont invariants et fourniront les mêmes π(s,i) par F. Définies seulement à un facteur multiplicatif près, les mesures γ et δ constituent une échelle de rapports. Une grosse différence, donc, entre les modèles M2 et M3 est que, lorsque F est spécifié, γ et δ sont des échelles ordinales en M2 tandis qu'elles sont de rapports en M3.

Remarquons qu'un modèle essentiellement équivalent à celui de (M3) serait issu de l'hypothèse que la probabilité de réussite ne dépend que de l'écart entre une compétence $\gamma'(s)$ et une difficulté $\delta'(i)$:

$$\pi\,(s,i) = \phi\{\gamma'(s) - \delta'(i)\} \qquad \text{(D3)}$$

où ϕ est une fonction croissante de R vers [0, 1], γ' une application de S vers R, δ' une application de I vers R.

Pour passer de la formulation (D3) à (M3) ou vice-versa, on utilise les transformations $\gamma'(s) = \ln(\gamma(s))$, $\delta'(i) = \ln(\delta(i))$ et leurs réciproques $\gamma(s) = \exp(\gamma'(s))$, $\delta(i)= \exp(\delta'(i))$.

Ainsi, $\pi(s,i) = \phi\{\ln(\gamma(s)) - \ln(\delta(i))\} = \phi\{\ln(\gamma(s)/\delta(i))\}$ devient (M3) avec $F = \phi(\ln)$; inversement $\pi(s,i) = F\{\exp(\gamma'(s))/\exp(\delta'(i))\} = F\{\exp(\gamma'(s)-\delta'(i))\}$ devient (D3) avec $\phi = F(\exp)$.

Les deux langages sont équivalents, le premier utilisant la structure multiplicative de R+, le second utilisant la structure additive de R. En prenant le logarithme népérien des compétences γ et difficultés δ reprises dans le Tableau 10, on obtient l'échelle des γ' et δ' :

	s_3	s_7	s_{10}	s_1	i_2	i_3	i_5	$\{i_6, s_2\}$	$\{i_1,s_4\}$	$\{i_4,s_5\}$	s_6	s_8	s_9
$\gamma(s)$ ou $\delta(i)$	0	2	6	8	9	10	11	12	13	14	16	18	19
$\gamma'(s)$ ou $\delta'(i)$	$-\infty$	0.69	1.79	2.08	2.20	2.30	2.40	2.48	2.56	2.64	2.77	2.89	2.94

C'est une échelle de différences car elle n'est définie, quand ϕ est spécifié, qu'à une constante additive près.

Il faut signaler enfin que pour formaliser la discrimination entre items, dès les années 20, Thurstone [1927a, 1927b] avait implicitement avancé le modèle (D3) en prenant pour ϕ la fonction de répartition d'une variable normale (pour une discussion complète, voir [Bock et Jones, 1968 ; Falmagne, 1985]).

LE MODÈLE DE RASCH

Le modèle le plus fort auquel nous arrivons, le modèle de Rasch, spécifie simplement la nature de la fonction F, la plus simple qui varie dans [0, 1], c'est-à-dire $F(x)=(1+x^{-1})^{-1}$:

$$\pi(s,i) = (1+\delta(i)/\gamma(s))^{-1} = \gamma(s)/(\gamma(s) + \delta(i)) \qquad \text{(M4)}$$

Puisque $1 - \pi(s,i)$, probabilité d'un échec, est $\delta(i)/(\gamma(s)+\delta(i))$, le modèle de Rasch a l'expression équivalente éclairante

$$\pi(s,i)/(1-\pi(s,i)) = \gamma(s)/\delta(i$$

qui s'énonce : la probabilité d'une réussite est à la probabilité d'un échec comme la compétence du sujet est à la difficulté de l'item [Rasch, 1960/1980 ; Fischer, 1995 ; voir aussi Wright et Mok, 2004].

De ceci découle que pour tous $s, t \in S$ et tous $i, j \in I$

$$\{\pi(s,i)/(1 - \pi(s,i))\} \times \{\pi(t,j)/(1-\pi(t,j))\} = \{\pi(s,j)/(1-\pi(s,j))\} \times \{\pi(t,i)/(1-\pi(t,i))\} \qquad \text{(C4)}$$

La preuve est immédiate : les deux membres de l'égalité valent

$$\{\gamma(s)/\delta(i)\} \times \{\gamma(t)/\delta(j)\} \text{ et } \{\gamma(s)/\delta(j)\} \times \{\gamma(t)/\delta(i)\}$$

On peut montrer [Andersen, 1973 ; Rasch, 1960 ; voir aussi Falmagne, 1985, p. 147-148 ; Roskam et Jansen, 1984] que (C4) est une condition suffisante de (M4). En d'autres termes, il existe des fonctions compétence γ et difficulté δ vérifiant (M4) si et seulement si toutes les probabilités de réussite satisfont à (C4).

Le modèle de Rasch renforce celui de (M3) et contraint plus les données d'observation. Vérifier que (C4) entraîne bien (C21), (C22) et (C3) sera un exercice léger si l'on remarque d'abord que l'inégalité $p<q$ est équivalente à l'inégalité $(p/(1-p)) < (q/(1-q))$ pour $p, q \in]0, 1[$.

Démontrons que (C4) implique (C21) : s'il existe $i \in I$ tel que $\pi(t,i) < \pi(s,i)$, on a $\{\pi(t,i)/(1-\pi(t,i))\} < \{\pi(s,i)/(1-\pi(s,i))\}$, donc

$$1 < \{\pi(s,i)(1-\pi(t,i))\} / \{(1-\pi(s,i))\pi(t,i)\}$$

Or cette dernière quantité vaut $\{\pi(s,j)(1-\pi(t,j))\}/\{(1-\pi(s,j))\pi(t,j)\}$ pour tout $j \in I$ par (C4). Ainsi $\{\pi(t,j)/(1-\pi(t,j))\} < \{\pi(s,j)/(1-\pi(s,j))\}$, donc $\pi(t,j) < \pi(s,j)$ pour tout $j \in I$.

La démonstration de (C22) est parallèle.

Démontrons enfin que (C4) implique (C3) : si $\pi(s,j) < \pi(t,k)$ et $\pi(t,i) < \pi(u,j)$, on a $\{\pi(s,j)/(1-\pi(s,j))\} < \{\pi(t,k)/(1-\pi(t,k))\}$ et $\{\pi(t,i)/(1-\pi(t,i))\} < \{\pi(u,j)/(1-\pi(u,j))\}$ qui donnent par multiplication

$$\{\pi(s,i)\, \pi(t,j)/(1 - \pi(s,i))(1-\pi(t,j))\} = \{\pi(s,j)\, \pi(t,i)/(1 - \pi(s,j))(1 - \pi(t,i))\}$$

$$<$$

$$\{\pi(t,k)\, \pi(u,j)/(1 - \pi(t,k))(1 - \pi(u,j))\} = \{\pi(t,j)\, \pi(u,k)/(1 - \pi(t,j))(1 - \pi(u,k))\}$$

En simplifiant aux extrêmes gauche et droite par $\{\pi(t,j)/(1-\pi(t,j))\}$, il reste

$$\{\pi(s,i)/(1 - \pi(s,i))\} < \{\pi(u,k)/(1 - \pi(u,k))\} \text{ c'est-à-dire } \pi(s,i) < \pi(u,k).$$

Contrastons les conditions (C21), (C22), (C3) qui ne font jouer que l'ordre des probabilités et la condition (C4) qui établit un lien algébrique entre les valeurs numériques de ces probabilités.

Que le modèle de Rasch définisse γ et δ comme échelle de rapports vient de ce qu'il particularise le modèle de (M3) et, par ailleurs, saute aux yeux en regardant sa formulation $(\pi(s,i)/(1-\pi(s,i)) = \gamma(s)/\delta(i)$. On voit le chemin parcouru depuis les modèles de (M1) et (M2) où γ et δ n'étaient que des échelles ordinales.

Arrivés à ce stade, les rapports entre les divers modèles peuvent se résumer dans le schéma repris dans le Tableau 11.

(M4) RASCH	⇒	(M3)	⇒	(M2)	⇒	(M1) GUTTMAN
⇕		⇕		⇕		⇕
(C4)	⇒	(C21), (C22), (C3)	⇒	(C21)	⇒	(C11)
						FERRERS
				(C22)	⇒	(C12)
				si $\pi(s,i) \geq \pi^*$	est identifié	à 's réussit i'

Tableau 11. Résumé des rapports entre les différents modèles (de Guttman à Rasch)

Ainsi des modèles de plus en plus forts sont introduits : Guttman / M2 / M3 / Rasch avec en parallèle des conditions de plus en plus astreignantes : Ferrers / C21 et C22/ C21, C22 et C3 / C4.

Il est décent de se poser la question : quel est l'intérêt de telles jongleries mathématiques? La réponse sort de l'opposition entre des modèles hypothétiques, tels ceux définis par (M1), (M2), (M3), (M4), qui sont des énoncés existentiels et leurs conditions nécessaires et suffisantes, telles (C11), (C12), (C21), (C22), (C3), (C4), qui sont des énoncés universels [Popper, 1973]. Les premiers disent que des observables directs (la relation 'réussir') ou indirects (les probabilités $\pi(s,i)$) sont liés à certains inobservables (γ, δ, g, F) ; ainsi exprimés ils ne peuvent être éprouvés. Les seconds, au contraire, sont testables car ils portent sur toutes les données. Voilà l'intérêt d'une formalisation du mesurage.

La fonction $(x, y) \rightarrow x/(x+y)$ n'est qu'un exemplaire très particulier parmi l'infinité des fonctions F croissantes en (x/y) postulées dans (M3). Le succès du modèle de Rasch vient moins d'une pertinence universelle contestable que de ses propriétés agréables dont la plus intéressante au point de vue de l'utilisation est qu'il permet une inférence statistique classique [voir Admane et Mesbah, 2006 ; Andersen, 1973 ; Andrich, 1988 ; Bond et Fox, 2007 ; Fischer, 1981, 2007 ; Flieller, 1994 ; Penta *et al.*, 2005].

INFÉRENCE STATISTIQUE

Supposons dorénavant que les observations résultent d'un échantillon au hasard de sujets et d'items. Pour illustrer la chose, reprenons la matrice du Tableau 6 des réussites/échecs qui manifestent presque une échelle de Guttman. Si ces données sont engendrées par un modèle de Rasch, comment estimer les compétences des 10 sujets et les difficultés des 6 items ?

Fischer [1981] a montré que les scores bruts des sujets et des items constituent des statistiques exhaustives pour les paramètres γ(s) et δ(i), la méthode usuelle d'estimation, dite du maximum de vraisemblance, l'ayant conduit aux équations

$$\text{score de s} = r(s) = \sum_i \frac{c(s)}{c(s)+d(i)} \qquad \text{score de i} = m(i) = \sum_s \frac{c(s)}{c(s)+d(i)} \tag{E}$$

c(s) (resp. d(i)) étant l'estimation de la compétence γ de s (resp. difficulté δ de i).

Fisher et Molenaar [1995] ont montré que cette propriété d'exhaustivité du score de s et de i permet de caractériser de façon unique le modèle de Rasch.

Les équations ci-dessus sont moins simples à résoudre et plus naturelles qu'il n'y paraît ; complexes d'abord parce qu'elles sont loin d'être linéaires en leurs inconnues ; néanmoins naturelles parce qu'elles identifient le score observé dans l'échantillon (r(s) ou m(i)) au score (de s ou i) attendu en théorie. En effet, bien que l'item i ne puisse apporter au score du sujet s que 1 (réussite) ou 0 (échec), en moyenne il apporterait π(s,i). Par conséquent, le score attendu de s est la somme des apports théoriques de chaque item, soit $\sum_i \pi(s,i)$.

De manière analogue, le score attendu de i est la somme des apports théoriques de chaque sujet, soit $\sum_s \pi(s,i)$. Compte tenu de (M4), les équations (E) sont obtenues en égalant les scores dans l'échantillon aux scores attendus.

Par (E), deux sujets de même score r(s) ont la même compétence estimée c(s) et deux items de même score m(i) ont la même difficulté estimée d(i). Cela entendu et avec le fait que $r(s_3) = 0$ demande $c(s_3) = 0$, l'estimation dans l'exemple du Tableau 6 consistera à résoudre

$$r(s) = \frac{c(s)}{c(s)+d(i_4)} + \frac{c(s)}{c(s)+d(i_1)} + \frac{c(s)}{c(s)+d(i_6)} + \frac{c(s)}{c(s)+d(i_5)} + \frac{2c(s)}{c(s)+d(i_3)}$$

Soient 5 équations que le lecteur complétera aisément :

$r(s_7) = 1 = \ldots$

$r(s_{10}) = r(s_1) = 2 = \ldots$

$r(s_2) = 3 = \ldots$

$r(s_4) = r(s_5) = r(s_6) = 4 = \ldots$

$r(s_8) = r(s_5) = 5 = \ldots$

simultanément à :

$$m(i) = \frac{c(s_7)}{c(s_7)+d(i)} + \frac{2c(s_{10})}{c(s_{10})+d(i)} + \frac{c(s_2)}{c(s_2)+d(i)} + \frac{3c(s_4)}{c(s_4)+d(i)} + \frac{2c(s_8)}{c(s_8)+d(i)}$$

Soient 5 autres équations

$m(i_4) = 1 = ...$

$m(i_1) = 2 = ...$

$m(i_6) = 4 = ...$

$m(i_5) = 7 = ...$

$m(i_3) = m(i_2) = 8 = ...$

pour en tirer les 10 inconnues $c(s_7)$, $c(s_{10}) = c(s_1)$, $c(s_2)$, $c(s_4) = c(s_5) = c(s_6)$, $c(s_8) = c(s_9)$, $d(i_4)$, $d(i_1)$, $d(i_6)$, $d(i_5)$, $d(i_3) = d(i_2)$.

En général, il y a autant d'équations (E) que d'inconnues c(s) et d(i), leur nombre commun étant la somme du nombre de scores distincts chez les sujets et du nombre de scores distincts chez les items, en omettant les scores nuls et parfaits [Fischer, 1981]. Des algorithmes itératifs doivent être exécutés sur ordinateur pour livrer les solutions. Ce recours à l'ordinateur est d'autant plus nécessaire que le nombre d'items et de sujets est assez élevé. Il existe aujourd'hui sur le marché de nombreux logiciels [Bond, Fox, 2007 ; Hambleton *et al.*, 1991 ; Linacre, 2006 ; Rizopoloulos, 2006].

Dans l'exemple, le tableau suivant montre en marges les c(s) et d(i) et, dans la matrice, les (arrondis à trois décimales des) probabilités estimées : à moins d'un centième près, les scores attendus coïncident avec les scores observés du Tableau 6.

		0	1.004	2.005	2.005	3.004	3.995	3.995	3.995	4.996	4.996	Scores attendus
57440	i_4	0	0	0	0	0.001	0.079	0.079	0.079	0.383	0.383	1.004
17540	i_1	0	0	0	0	0.003	0.220	0.220	0.220	0.670	0.670	2.003
2170	i_6	0	0	0	0	0.026	0.696	0.696	0.696	0.943	0.943	4.000
1.07	i_5	0	0.140	0.437	0.437	0.982	1	1	1	1	1	6.996
0.2292	i_3	0	0.432	0.784	0.784	0.996	1	1	1	1	1	7.996
0.2292	i_2	0	0.432	0.784	0.784	0.996	1	1	1	1	1	7.996
d(i)		s_3	s_7	s_{10}	s_1	s_2	s_4	s_5	s_6	s_8	s_9	
		0	0.1744	0.83	0.83	58.5	4960	4960	4960	35600	35600	c(s)

Tableau 12. estimation des γ(s) et δ(i) à partir des scores bruts, détermination des probabilités de réussite et des scores attendus

Finalement, le mesurage conjoint des sujets et items se présente dans l'ordre :

	s_3	s_7	$\{i_2, i_3\}$	$\{s_{10}, s_1\}$	i_5	s_2	i_6	$\{s_4, s_5, s_6\}$	i_1	$\{s_8, s_9\}$	i_4
c(s) ou d(i)	0	0.1744	0.2292	0.83	1.07	58.5	2170	4960	17540	35600	57440

Tableau 13. Positionnement des sujets et des items sur l'échelle de rapports de Rasch

En traduisant l'échelle de rapports en échelle de différences comme définie par (D3), le modèle (M4) devient

$$\pi(s,i) = 1/(1 + \exp(\delta(i) - \gamma(s))) = \exp(\gamma(s)) / (\exp(\gamma(s)) + \exp(\delta(i)))$$

C'est la présentation la plus connue du modèle de Rasch où la probabilité de réussite est fonction logistique $z \to 1/(1+\exp(-z))$ de l'écart $\gamma(s) - \delta(i)$ entre des mesures de la compétence et de la difficulté. L'appellation modèle logistique (simple) est justifiée dans ce cas.

En prenant le logarithme népérien des c(s) et d(i) précédents, on obtient une échelle de différences

s_3	s_7	$\{i_2, i_3\}$	$\{s_{10}, s_1\}$	i_5	s_2	i_6	$\{s_4, s_5, s_6\}$	i_1	$\{s_8, s_9\}$	i_4
$-\infty$	-1.746	-1.473	-0.186	0.068	4.069	7.682	8.509	9.772	10.480	10.958

Tableau 14. Positionnement des sujets et des items
sur l'échelle des différences de Rasch

Il n'est pas douteux que les gens sont plus familiers de la lecture de cette graduation, parce qu'elle est un type particulier d'échelles d'intervalles. Une autre est obtenue en y ajoutant par exemple 4.605 = ln 100, ce qui revient à centupler les valeurs de l'échelle de rapports.

Quatre cas limites des équations (E) posent problème quand tous les termes des sommes $\sum_i$ ou $\sum_s$ doivent être 0 ou 1.

r(s) = 0 exige c(s)=0,

r(s) = « nombre d'items présents » exige $\frac{c(s)}{c(s)+d(i)} = 1$ pour tout i, soit $c(s) = \infty$,

m(i) = 0 exige $\frac{c(s)}{c(s)+d(i)} = 0$ pour tout s, soit $d(i) = \infty$,

m(i) = « nombre de sujets présents » exige d(i) = 0.

Ces phénomènes ne sont pas dérangeants si l'on songe qu'il ne s'agit que d'estimations des paramètres $\gamma(s)$ et $\delta(i)$ qui eux ne sortent pas de R^+. Par exemple, $c(s) = \infty$ ne signifie qu'une intensité extrême de la compétence ; $\gamma(s)$ est bien fini, mais très grand, en tout cas supérieur à tout autre paramètre, ce qui fait que s réussit tous les items. Ou bien, d(i) = 0 manifeste une facilité excessive, $\delta(i)$ est encore positif mais très petit, en tout cas inférieur à tout autre paramètre, de telle manière que i soit réussi par n'importe quel sujet, aussi petite soit sa compétence.

Enfin il faut signaler que d'autres méthodes d'estimation sont plus appropriées dans certaines conditions d'échantillonnage : items (resp. sujets) considérés pour eux-mêmes, et non comme représentants aléatoires d'un univers d'items (resp. de sujets) ; compétences des sujets (resp. difficultés des items) déjà connues par ailleurs ; etc. Dans tous les cas, il est prouvé que la statistique r(s) (resp. m(i)) est exhaustive, au sens technique du terme, pour estimer les $\gamma(s)$ (resp. $\delta(i)$) [Admane, Mesbah, 2006 ; Doran *et al*, 2007 ; Feddag, Mesbah, 2006]. Cette propriété, autre caractéristique du modèle de Rasch, explique aussi en grande partie l'utilisation et le succès de ce modèle de mesure.

CONCLUSION

Lorsque la réussite ou l'échec ne dépend que d'une certaine compétence γ mesurable chez les sujets et d'une certaine difficulté δ mesurable chez les items, le modèle le plus simple est

M1 : le sujet s réussit l'item i si et seulement si $\gamma(s) \geq \delta(i)$.

Dans le cas d'ensembles finis de sujets et d'items, on prouve que ce modèle tient si et seulement si l'une (C11) ou l'autre (C12) version de la condition de Ferrers est vérifiée.

C11 : s'il existe un item réussi par le sujet s et raté par le sujet t, alors tout item raté par s est raté par t.

C12 : s'il existe un sujet qui réussit l'item j et rate l'item i, alors tout sujet qui réussit i réussit j.

La démonstration passe par la création d'une relation β dans la réunion des ensembles de sujets et d'items qui exprime de manière naturelle la dominance d'un sujet / item sur un sujet / item. Par sa définition, β est réflexive, transitive et totale, d'où il résulte mathématiquement que sa partie symétrique est une équivalence et sa partie antisymétrique est un ordre strict total. Les classes d'équivalence sont alors totalement ordonnées et peuvent donc se succéder intrinsèquement en une seule série. Il suffit d'associer à cette liste une suite croissante quelconque de nombres pour construire de bonnes mesures γ et δ. Définies à une transformation croissante près, γ et δ sont des échelles ordinales conjointes. Le modèle M1 est apparu dans les sciences humaines sous le nom d'échelle de Guttman.

Dans la réalité, compétence des sujets et difficulté des items ne jouent pas seules pour engendrer réussite ou échec. S'introduisent généralement de légères perturbations dues à d'autres causes inconnues ou délibérément ignorées. Cette composante aléatoire est prise en compte en attribuant à telle compétence confrontée à telle difficulté seulement une probabilité de réussite :

$$M2 : \pi(s,i) = g(\gamma(s), \delta(i)).$$

Pour que ce deuxième modèle soit cohérent, la fonction g de R^2 vers [0, 1] doit être croissante en $\gamma(s)$ et décroissante en $\delta(i)$. On prouve qu'il existe un tel g accompagné des mesures γ et δ si et seulement si les deux conditions C21 et C22 portant sur les $\pi(s,i)$ sont satisfaites.

C21 : s'il existe un item i avec $\pi(t,i) < \pi(s,i)$, alors $\pi(t,j) \leq \pi(s,j)$ pour tout item j

C22 : s'il existe un sujet s avec $\pi(s,i) < \pi(s,j)$, alors $\pi(t,i) \leq \pi(t,j)$ pour tout sujet t

Ce modèle M2 n'est qu'une version assouplie de l'échelle de Guttman car, d'une part, C21 et C22 sont des formes probabilistes de C11 et C12, d'autre part, si on y fige un seuil π^* et si on définit conventionnellement la réussite de s sur i par $\pi(s,i) \geq \pi^*$, un modèle M1 est obtenu quelle que soit la probabilité critique π^* entre 0 et 1.

Le modèle M3 renforce le deuxième en demandant que la probabilité de réussite ne dépende que du rapport entre compétence et difficulté :

$$M3 : \pi(s, i) = F(\gamma(s)/\delta(i))$$

où F est une fonction croissante. Puisque M3 particularise M2, il est logique qu'une condition nécessaire et suffisante du modèle M3 consiste à adjoindre à C21 et C22 une autre condition C3.

$$\text{C3 : pour tous } s, t, u \in S \text{ et tous } i, j, k \in I$$

$$\pi(s,j) < \pi(t,k) \text{ et } \pi(t,i) < \pi(u,j) \text{ entraînent } \pi(s,i) \leq \pi(u,k)$$

Une grosse différence entre les modèles M2 et M3 est que, lorsque F est spécifié, γ et δ sont des échelles ordinales en M2 tandis qu'elles sont de rapports en M3.

Enfin, quand on choisit pour F la fonction la plus simple qui varie dans [0, 1], on obtient le modèle

$$\text{M4 : } \pi(s,i) = \gamma(s)/(\gamma(s) + \delta(i))$$

Celui-ci est caractérisé par l'unique condition C4

$$\text{C4 : pour tous } s, t \in S \text{ et tous } i, j \in I$$

$$\{\pi(s,i)/(1 - \pi(s,i))\} \times \{\pi(t,j)/(1-\pi(t,j))\} = \{\pi(s,j)/(1-\pi(s,j))\} \times \{\pi(t,i)/(1-\pi(t,i))\}$$

laquelle, évidemment, entraîne C21, C22 et C3.

Quand les échelles de rapports de la compétence et de la difficulté sont transformées en échelles de différences par logarithme, le modèle de Rasch se présente sous la forme logistique

$$\text{Modèle de Rasch : } \pi(s, i) = 1/(1 + \exp(\delta(i) - \gamma(s)))$$

qui est la plus connue des utilisateurs. Le modèle de Rasch a l'intérêt supplémentaire de s'insérer dans l'inférence statistique. Avec un échantillon au hasard, les compétences des sujets et les difficultés des items peuvent être estimées à partir de leurs scores par une procédure itérative.

Pour nous résumer, des modèles de plus en plus forts sont introduits : Guttman / M2 / M3 / Rasch avec en parallèle des conditions de en plus astreignantes : Ferrers / C21 et C22 / C21, C22 et C3 / C4. Ces conditions ont l'avantage décisif sur la définition des modèles – qui font jouer des inobservables – de porter sur toutes les données et, par conséquent, de pouvoir être testées dans des observations.

Cette leçon de logique constructive voulait montrer qu'il est profitable de fonder toute mesure. Le lecteur du présent article consultera avec profit les notes bibliographiques sur la théorie du mesurage des travaux de Falmagne [1988, 1994], bien que ces travaux nécessitent une certaine maturité mathématique pour être complètement exploités. Espérons dès à présent que notre lecteur ne sera pas de ceux qui se satisfont du ronron statistique, s'évertuant à tirer des conclusions de mesures insignifiantes au sens propre du mot.

Remerciements. Nous tenons à remercier vivement le professeur Claude Flament qui nous a encouragés à publier ce texte de synthèse, ainsi que les étudiants des universités de Liège (Belgique) et d'Aix-Marseille I (France), qui ont contribué – par leurs questions critiques – à améliorer les aspects pédagogiques du texte lors des leçons de psychologie mathématique dispensés par Ch. Heuchenne et A-.S. El Ahmadi entre 1985 et 1994 à Liège et par A.-S. El Ahmadi et D. Bertrand depuis 1997 à Aix-en-Provence. Nos vifs remerciements enfin aux deux experts pour les remarques judicieuses dont ils nous ont fait part lors de l'examen du manuscrit.

BIBLIOGRAPHIE

ADMANE O., MESBAH M., « Estimation des paramètres du modèle de Rasch dichotomique », *Revue de l'Institut de Statistique de l'Université de Paris* 50, 2006, p. 3-30.

ANDERSEN E. B., "A goodness of fit test for the Rasch model", *Psychometrika* 38, 1973, p. 123-140.

ANDRICH D., "An elaboration of Guttman scaling with Rasch models for measurement", N. Brandon-Tuma (ed.), *Sociological Methodology*, San Francisco, Jossey-Bass., Chapter 2, 1985, p. 33-80.

ANDRICH D., *Rasch models for measurement*, Sage University Paper series on Quantitative Applications in the Social Sciences, series n° 07-068, Beverly Hills, Sage Publications, 1988.

BIRKHOFF G., *Lattice Theory*, Providence (Rhode Island), American Mathematical Society Colloquium Publication N° XXV, 1967.

BOCK R. D., JONES L. V., *The measurement and prediction of judgement and choice*, San Francisco, Holden-Day, 1968.

BOND. T. G., FOX C. M., *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*, 2nd edition, Mahwah (New Jersey), Lawrence Erlbaum Associates, 2007.

DOIGNON J.-P., DUCAMP A., FALMAGNE J.-C., "On realizable biorders and the biorder dimension of a relation", *Journal of Mathematical Psychology* 28, 1984, p. 73-109.

DORAN H., BATES D., BLIESE P., DOWLING M., "Estimating the multilevel Rasch model: with lme4 package", *Journal of Statistical Software* 20, 2007, [http://www. jstatsoft.org/v20/i02/paper].

DUCAMP A., FALMAGNE J.-C., "Composite measurement", *Journal of Mathematical Psychology* 6, 1969, p. 359-390.

FALMAGNE J.-C., *Elements of Psychophysical Theory*, Oxford, Clarendon Press, 1985.

FALMAGNE J.-C., « Propos sur la théorie du mesurage », *Mathématiques, Informatique et Sciences humaines* 101, 1988, p. 7-34.

FALMAGNE J.-C., « Mesurage, modèles mathématiques et psychophysique », M. Richelle, J. Requin et M. Robert (éds), *Traité de Psychologie Expérimentale,* Vol.1, Paris, Presses Universitaires de France, 1994 p.79-141.

FEDDAG M. L., MESBAH M., "Approximate estimation in generalized linear mixed models with applications to the Rasch model", *Computers and Mathematics with Applications* 51, 2006, p. 269-278.

FISCHER G. H., "On the existence and uniqueness of maximum likelihood estimates of the Rasch model" *Psychometrika* 46, 1981, p. 59-77.

FISCHER G. H., "Derivations of the Rasch Model", G.H. Fischer & I.W. Molenaar (Eds), *Rasch Models: Foundations, Recent Developments, and Applications*, New-York (NY), Springer-Verlag, 1995, p. 15-38.

FISCHER G. H., "Rasch Models", C. R. Rao & S. Sinharay (eds), *Handbook of Statistics,* vol. 26 *Psychometrics*, Amsterdam (The Netherlands), 2007, p. 515-586.

FISCHER G. H., MOLENAAR I. W., *Rasch Models: Foundations, Recent Developments, and Applications.* New-York (NY), Springer-Verlag, 1995.

FLIELLER A., « Méthodes d'étude de l'adéquation au modèle logistique à un paramètre (modèle de Rasch) », *Mathématiques, Informatique et Sciences humaines* 127, 1994, p. 19-47.

GUTTMAN L., "A basis for scaling qualitative data", *American Sociological Review* 9, 1944, p. 139-150.

HAMBLETON R., SWAMINATHAN H., ROGERS H., *Fundamentals of item response theory*, London, Sage Publications Ltd, 1991.

KRANTZ D. H., LUCE R., SUPPES P., TVERSKY A., *Foundations of Measurement,* Vol. 1: *Additive and Polynomial Representations*, New York (N.Y.), Academic Press, 1971.

LINACRE J.-M., "Stochastic Guttman order", *Rasch Measurements Transactions* 5(4), 189, 1992.

LINACRE J.-M., "A user's guide and manual to Winsteps and Ministeps Rasch-model computer programs. Chicago", 2006, [http://www.winsteps.com].

MONJARDET B., « Axiomatiques et propriétés des quasi-ordres », *Mathématiques et Sciences humaines* 63, 1978, p. 51-82.

PENTA M., ARNOULD C., DECRUYNAERE C., *Développer et interpréter une échelle de mesure : applications du modèle de Rasch*, Sprimont, Mardaga, 2005.

POPPER K., *La logique de la découverte scientifique*, Paris, Payot, 1973. ;

RASCH G., *Probabilistic Models for some Intelligence and Attainment Tests*, Copenhaguen, The Danish Institute of Educational Research, 1960 [reprinted by the Chicago University Press, 1980].

RIGUET J., « Les relations de Ferrers », *Comptes rendus de l'Académie des Sciences de Paris* 232, 1951, p. 1729-1730.

RIZOPOLOULOS D., "ltm: An R package for latent variable modeling and Item Response Analyses", *Journal of Statistical Software* 17, 2006, [http://www.jstatsoft.org/v17/i05/].

ROSKAM E. E., JANSEN P. G. W., "A new derivation of the Rasch model", E. Degreef & J. Van Buggenhaut (eds), *Trends in Mathematical Psychology*, Amsterdam , Elsevier Science Publishers, 1984, p. 293-307.

STOUFFER S. A., GUTTMAN L., SCHUMAN E. A., LAZARSFELD D., STAR S. A. CLAUSEN J. A., *Measurement and prediction*, Princeton (New Jersey), Princeton University Press, 1950.

SUPPES P., ZINNES J.-L., "Basic measurement theory", R.D. Luce, R.R. Bush & E. Galanter (eds), *Handbook of Mathematical Psychology*, Vol.1, New York, Wiley, 1963, p. 1-76.

THURSTONE L. L., "A law of comparative judgement", *Psychophysical Review* 34, 1927(a), p. 273-286.

THURSTONE L. L., "Psychological analysis", *American Journal of Psychology* 38, 1927(b), p. 368-389.

WRIGHT B., MOK M. M. C., "An overview of the family of Rasch measurement models", E. Smith & R. Smith (eds), *Introduction to Rasch Measurement: Theory, Models, and Applications*, Maple Grove, MN, JAM Press, 2004, p. 1-24.