

Avril 2004

Numéro 35

# Cahiers de l'IMA

Comment s'assurer de l'alignement  
d'un ensemble d'items

**Jean-Philippe Antonietti**

Institut de Mathématiques Appliquées  
Faculté des S.S.P.  
Université de Lausanne  
BFSH 2  
1015 Lausanne

# Comment s'assurer de l'alignement d'un ensemble d'items

Jean-Philippe Antonietti

## Introduction

Ces dernières décennies, une véritable révolution a eu lieu dans le champ de la mesure psychologique et éducationnelle. La théorie classique des scores s'est fait supplanter par la théorie de la réponse à l'item [4]. Cette nouvelle théorie est, entre autres, à l'origine de la création de vastes banques d'items et du développement des tests taillés sur mesure.

Les modèles de la réponse à l'item ne sont toutefois pas la panacée. Leur application nécessite que de nombreuses hypothèses fort astreignantes soient satisfaites. C'est pour pallier à ce défaut que d'autres modèles moins contraignants ont été proposés. Nous avons décrit dans un cahier précédent [1] l'un de ces modèles : le modèle de la réponse à l'item non-paramétrique de Mokken [7, 8, 9].

Pour que ce modèle soit applicable, quatre hypothèses doivent être satisfaites : l'unidimensionnalité, l'indépendance locale, la monotonie des courbes caractéristiques des items et la non-intersection des courbes caractéristiques des items. Afin d'illustrer notre propos, donnons un exemple.

Supposons que l'on veuille mesurer les compétences mathématiques d'élèves de 4<sup>e</sup> année primaire en recourant au modèle de Mokken. Quatre conditions doivent être satisfaites. La première impose que tous les items de l'épreuve de mathématiques soumise aux élèves évaluent un et un seul trait latent : les compétences mathématiques. Les réponses des élèves ne devraient donc pas être influencées par leur niveau en lecture. Les compétences mathématiques devraient être homogènes : il ne devrait pas y avoir de différences entre, par exemple, compétences numériques et compétences spatiales. La deuxième condition impose que les réponses d'un élève ne dépendent que de ses compétences et non pas de l'ordre dans lequel les questions lui sont posées. La troisième condition exige que, si deux élèves ne sont pas aussi compétent l'un que l'autre, le plus compétent ait plus de chance de répondre correctement à n'importe quelle question que celui qui l'est moins. La quatrième et dernière condition impose que l'ordre de difficulté des items ne dépende pas des compétences des élèves.

Pratiquement ces conditions ne peuvent être vérifiées qu'a posteriori. Mais

comment? C'est ce que nous allons montrer dans ce travail en mettant plus particulièrement l'accent sur la première condition du modèle, l'hypothèse d'unidimensionnalité.

Dans le cadre d'un modèle de la réponse à l'item paramétrique, plusieurs techniques permettent de tester l'hypothèse d'unidimensionnalité [2, 3, 4, 5, 6]. Décrivons l'une de ces méthodes qui s'appuie sur les principes de l'analyse parallèle<sup>1</sup>. Les étapes d'une telle analyse sont les suivantes :

1. Effectuer une analyse factorielle à partir de la matrice de corrélation des items du test. Calculer les valeurs propres associées.
2. Générer un jeu de données parallèle unidimensionnel selon le modèle de la réponse à l'item. Les sujets du jeu de données parallèle occupent le long du trait latent les mêmes positions que les sujets observés. Les items du jeu parallèle possèdent les mêmes caractéristiques que les items réels (*i.e.* même difficulté, même pouvoir discriminant et même devinement).
3. Effectuer une analyse factorielle du jeu de données parallèle. Calculer les valeurs propres associées.
4. Comparer la deuxième valeur propre du jeu réel à celle du jeu parallèle. Si la valeur empirique est suffisamment proche de la valeur théorique obtenue à partir de l'analyse du jeu parallèle (*i.e.* leur rapport est suffisamment prêt de 1) le jeu de donnée réel est considéré comme unidimensionnel<sup>2</sup>.

Cette méthode n'est pas applicable aux modèles plus généraux de la réponse à l'item non-paramétrique. Certains outils existent néanmoins. À cette fin, Mokken [7] propose d'utiliser différents coefficients d'alignement (*scalability*). L'un de ces coefficients permet de mesurer l'alignement entre deux items, un autre permet de mesurer l'alignement d'un item avec les autres items du test et un dernier permet de mesurer l'alignement global d'un ensemble d'items. Nous allons dans la section suivante définir ces coefficients et décrire quelques-unes de leurs propriétés.

---

1. Rappelons que l'analyse parallèle s'utilise pour déterminer la dimensionnalité d'un jeu de données.

2. Au lieu de ne générer qu'un jeu de données parallèle, il est possible d'en générer 1000 par exemple, d'effectuer l'analyse factorielle de chacun, de calculer les valeurs propres associées puis de déterminer la valeur du 95<sup>e</sup> centile de la distribution des deuxièmes valeurs propres  $\lambda_{crit}^{(2)}$ . Une règle de décision mieux fondée consiste alors à considérer que le jeu réel n'est pas unidimensionnel si la deuxième valeur propre empirique  $\lambda_{emp}^{(2)}$  est supérieure à la valeur propre critique  $\lambda_{crit}^{(2)}$ .

# 1 Coefficients d'alignement

## 1.1 Définitions

Considérons deux items dichotomiques  $X_i$  et  $X_j$ . Supposons que l'item  $i$  soit plus difficile que l'item  $j$ . Ainsi  $P_i = P(X_i = 1) < P_j = P(X_j = 1)$ . Supposons aussi que la probabilité de répondre correctement aux deux items soit égale à  $P_{ij}$  :

$$P_{ij} = P(X_i = 1, X_j = 1). \quad (1)$$

Les probabilités associées à chaque événement possible sont rapportées dans le Tableau 1.

	$X_i = 0$	$X_i = 1$	Total
$X_j = 0$	$1 - P_i - P_j + P_{ij}$	$P_i - P_{ij}$	$1 - P_j$
$X_j = 1$	$P_j - P_{ij}$	$P_{ij}$	$P_j$
Total	$1 - P_i$	$P_i$	1

TABEAU 1 – Valeur des probabilités dans le cas général.

Dans cette situation, la covariance entre  $X_i$  et  $X_j$  est égale à :

$$\text{cov}(X_i, X_j) = P_{ij} - P_i P_j. \quad (2)$$

Si le modèle était déterministe (modèle de Guttman), la probabilité qu'un sujet réponde correctement à la question  $i$  mais se trompe à la question  $j$  serait égale à zéro. Les probabilités des différents événements correspondant à cette situation sont définies dans le Tableau 2.

	$X_i = 0$	$X_i = 1$	Total
$X_j = 0$	$1 - P_j$	0	$1 - P_j$
$X_j = 1$	$P_j - P_i$	$P_i$	$P_j$
Total	$1 - P_i$	$P_i$	1

TABEAU 2 – Valeur des probabilités dans le cas d'un modèle déterministe.

Dans ce cas, les probabilités marginales étant fixées, la covariance entre  $X_i$  et  $X_j$  serait maximale; elle vaudrait :

$$\text{cov}_{\max}(X_i, X_j) = P_i - P_i P_j = P_i(1 - P_j). \quad (3)$$

Le coefficient d'alignement  $H_{ij}$  entre l'item  $i$  et l'item  $j$  est défini comme le rapport entre  $\text{cov}(X_i, X_j)$  et  $\text{cov}_{\max}(X_i, X_j)$  :

$$H_{ij} = \frac{\text{cov}(X_i, X_j)}{\text{cov}_{\max}(X_i, X_j)} = \frac{P_{ij} - P_i P_j}{P_i - P_i P_j} = 1 - \frac{P_i - P_{ij}}{P_i(1 - P_j)}. \quad (4)$$

Dans le cas déterministe  $P_{ij} = P_i$ ,  $H_{ij}$  est donc égal à 1. Les courbes caractéristiques des items sont des fonctions escalier. En cas d'indépendance  $P_{ij} = P_i P_j$  et  $H_{ij}$  vaut donc 0. L'une des courbes caractéristiques au moins est non discriminante (*i.e.* horizontale).

De manière similaire, le coefficient d'alignement  $H_i$  entre l'item  $i$  et le reste des items  $R(i) = \sum_{k \neq i} X_k$  est défini comme le rapport entre la covariance observée  $\text{cov}(X_i, R(i))$  et la covariance maximale  $\text{cov}_{\max}(X_i, R(i))$  obtenue dans le cas d'un modèle déterministe<sup>3</sup> :

$$\begin{aligned} H_i &= \frac{\text{cov}(X_i, R(i))}{\text{cov}_{\max}(X_i, R(i))} = \\ &= \frac{\sum_{k \neq i} (P_{ik} - P_i P_k)}{\sum_{k > i} (P_i - P_i P_k) + \sum_{k < i} (P_k - P_i P_k)}. \end{aligned} \quad (5)$$

Finalement, le coefficient  $H$  permettant d'évaluer l'alignement global de l'ensemble des items d'un test se calcule de manière analogue en divisant la somme des covariances entre chaque item et le reste du test et la somme des covariances maximales associées :

$$\begin{aligned} H &= \frac{\sum_i \text{cov}(X_i, R(i))}{\sum_i \text{cov}_{\max}(X_i, R(i))} = \\ &= \frac{\sum_i \sum_{k \neq i} (P_{ik} - P_i P_k)}{\sum_i \sum_{k > i} (P_i - P_i P_k) + \sum_i \sum_{k < i} (P_k - P_i P_k)}. \end{aligned} \quad (6)$$

Les coefficients d'alignement théoriques s'estiment à partir des observations faites sur un échantillon en substituant dans les Équations 4, 5 et 6 les probabilités par leurs estimations. Ainsi le coefficient d'alignement entre les items  $i$  et  $j$  s'estime par  $\hat{H}_{ij}$  :

$$\hat{H}_{ij} = \frac{f_{ij} - f_i f_j}{f_i - f_i f_j} = 1 - \frac{f_i - f_{ij}}{f_i(1 - f_j)}, \quad (7)$$

avec

$f_i$  : la fréquence de réussite de l'item  $i$  dans l'échantillon;

$f_j$  : la fréquence de réussite de l'item  $j$  dans l'échantillon;

---

3. Par commodité nous supposons que les items sont ordonnés selon leur facilité, c'est-à-dire que si  $k > i$ , alors  $P_k > P_i$  et que si  $k < i$ , alors  $P_k < P_i$ .

$f_{ij}$  : la fréquence conjointe de la réussite de l'item  $i$  et de l'item  $j$  dans l'échantillon.

Le coefficient d'alignement entre l'item  $i$  et le reste des items s'estime par  $\hat{H}_i$ :

$$\hat{H}_i = \frac{\sum_{k \neq i} (f_{ik} - f_i f_k)}{\sum_{k > i} (f_i - f_i f_k) + \sum_{k < i} (f_k - f_i f_k)}. \quad (8)$$

Et, de manière similaire, le coefficient d'alignement global s'estime par  $\hat{H}$  :

$$\hat{H} = \frac{\sum_i \sum_{k \neq i} (f_{ik} - f_i f_k)}{\sum_i \sum_{k > i} (f_i - f_i f_k) + \sum_i \sum_{k < i} (f_k - f_i f_k)}. \quad (9)$$

## 1.2 Propriétés

Voici quelques propriétés que satisfont les coefficients l'alignement.

PROPRIÉTÉ 1

$$\min_k H_{ik} \leq H_i \leq \max_k H_{ik}.$$

**Démonstration.** Soient deux items quelconques  $i$  et  $k$  ( $i \neq k$ ). Nommons  $P_{ik}^0$  la probabilité de répondre correctement à l'item le plus difficile et de se tromper à l'item le plus simple en cas d'indépendance :

$$P_{ik}^0 = \begin{cases} P_i(1 - P_k) & \text{si } i < k \\ (1 - P_i)P_k & \text{si } i > k \end{cases} \quad (10)$$

Le coefficient d'alignement entre  $i$  et le reste du test peut s'exprimer comme une moyenne pondérée des coefficients d'alignement  $H_{ik}$  calculés entre l'item  $i$  et les autres items du test :

$$H_i = \frac{\sum_{k \neq i} P_{ik}^0 H_{ik}}{\sum_{k \neq i} P_{ik}^0}. \quad (11)$$

Nommons  $m_i$  le plus petit coefficient  $H_{ik}$  et  $M_i$  le plus grand.  $\min_k H_{ik} \leq H_i \leq \max_k H_{ik}$ , en d'autres termes  $m_i \leq H_i \leq M_i$ . Il s'ensuit que :

$$\frac{\sum_{k \neq i} P_{ik}^0 m_i}{\sum_{k \neq i} P_{ik}^0} \leq \frac{\sum_{k \neq i} P_{ik}^0 H_{ik}}{\sum_{k \neq i} P_{ik}^0} \leq \frac{\sum_{k \neq i} P_{ik}^0 M_i}{\sum_{k \neq i} P_{ik}^0}. \quad (12)$$

$m_i$  et  $M_i$  sont des termes constants qui peuvent être mis en évidence. Après simplification l'équation précédente devient :

$$\min_k H_{ik} \leq \frac{\sum_{k \neq i} P_{ik}^0 H_{ik}}{\sum_{k \neq i} P_{ik}^0} \leq \max_k H_{ik} \quad (13)$$

■

PROPRIÉTÉ 2

$$\min_i H_i \leq H \leq \max_i H_i.$$

**Démonstration.** Le coefficient d'alignement global peut s'exprimer comme une moyenne pondérée des coefficients d'alignement  $H_{ik}$  ou  $H_i$  :

$$H = \frac{\sum_i \sum_{k \neq i} P_{ik}^0 H_{ik}}{\sum_i \sum_{k \neq i} P_{ik}^0} = \frac{\sum_i \left( \sum_{k \neq i} P_{ik}^0 \right) H_i}{\sum_i \left( \sum_{k \neq i} P_{ik}^0 \right)} \quad (14)$$

Pour les mêmes raisons que précédemment, il en découle que

$$\min_{i, k} H_{ik} \leq H \leq \max_{i, k} H_{ik} \quad (15)$$

et que

$$\min_i H_i \leq H \leq \max_i H_i. \quad (16)$$

■

Avant d'énoncer encore d'autres propriétés importantes des coefficients d'alignement, nous allons démontrer que si les hypothèses d'unidimensionnalité, d'indépendance locale et de monotonie sont satisfaites, alors la covariance de n'importe quelle paire d'items est, dans la population, toujours positive ou nulle [11, p. 155].

**Démonstration.** Supposons que la distribution des capacités  $\theta$  au sein de la population soit définie de  $-\infty$  à  $+\infty$  par la densité de probabilité  $g(\theta)$ . Cela signifie que :

$$(\forall \theta) g(\theta) \geq 0 \quad \text{et} \quad \int_{-\infty}^{+\infty} g(\theta) d\theta = 1. \quad (17)$$

La popularité globale de l'item  $i$  vaut donc<sup>4</sup> :

$$P_i = P(X_i = 1) = \int P_i(\theta) g(\theta) d\theta. \quad (18)$$

En vertu de l'indépendance locale, la probabilité de résoudre correctement les items  $i$  et  $j$  vaut :

$$P_{ij}(\theta) = P(X_i = 1, X_j = 1 \mid \theta) = P_i(\theta) \cdot P_j(\theta). \quad (19)$$

Ainsi, globalement, la probabilité d'obtenir le score 1 aux items  $i$  et  $j$  vaut :

$$P_{ij} = P(X_i = 1, X_j = 1) = \int P_i(\theta) P_j(\theta) g(\theta) d\theta. \quad (20)$$

Les items ne pouvant prendre que les valeurs 0 ou 1, la covariance de  $X_i$  et  $X_j$  s'exprime ainsi :

$$\text{cov}(X_i, X_j) = P_{ij} - P_i P_j. \quad (21)$$

Dans l'Équation 21, substituons  $P_{ij}$ ,  $P_i$  et  $P_j$  par leur valeur et multiplions  $P_{ij}$  par  $\int g(\theta) d\theta$  qui est égal à 1. L'expression de la covariance de  $X_i$  et  $X_j$  devient :

$$\begin{aligned} \text{cov}(X_i, X_j) &= \int g(\theta) d\theta \times \int P_i(\theta) P_j(\theta) g(\theta) d\theta \\ &\quad - \int P_i(\theta) g(\theta) d\theta \times \int P_j(\theta) g(\theta) d\theta. \end{aligned} \quad (22)$$

Pour pouvoir ultérieurement effectuer des simplifications, partageons chaque terme du deuxième membre de l'Équation 22 en deux parties égales et utilisons la propriété selon laquelle la valeur d'une intégrale ne dépend pas du nom de la variable d'intégration. Ainsi :

$$\begin{aligned} \text{cov}(X_i, X_j) &= \frac{1}{2} \int g(\eta) d\eta \times \int P_i(\theta) P_j(\theta) g(\theta) d\theta \\ &\quad + \frac{1}{2} \int g(\theta) d\theta \times \int P_i(\eta) P_j(\eta) g(\eta) d\eta \end{aligned}$$

---

4. Dorénavant, pour alléger l'écriture, nous omettrons les bornes d'intégration qui seront toujours  $-\infty$  et  $+\infty$ .



$$\begin{aligned}
& -\frac{1}{2} \int P_i(\theta) g(\theta) d\theta \times \int P_j(\eta) g(\eta) d\eta \\
& -\frac{1}{2} \int P_i(\eta) g(\eta) d\eta \times \int P_j(\theta) g(\theta) d\theta \quad (23)
\end{aligned}$$

En regroupant tous les termes sous une seule intégrale double, nous obtenons :

$$\begin{aligned}
\text{cov}(X_i, X_j) &= \\
&= \frac{1}{2} \int \int [P_i(\theta) P_j(\theta) - P_i(\theta) P_j(\eta) \\
&\quad + P_i(\eta) P_j(\eta) - P_i(\eta) P_j(\theta)] g(\theta) g(\eta) d\theta d\eta \\
&= \frac{1}{2} \int \int [(P_i(\theta) - P_i(\eta)) \cdot (P_j(\theta) - P_j(\eta))] g(\theta) g(\eta) d\theta d\eta. \quad (24)
\end{aligned}$$

Le plan  $(\theta, \eta)$  peut être partagé en deux parties. Dans la première  $\theta > \eta$  et dans la seconde  $\theta \leq \eta$ .

En vertu de l'hypothèse de monotonie si  $\theta > \eta$ , alors  $P_i(\theta) \geq P_i(\eta)$  et  $P_j(\theta) \geq P_j(\eta)$ . Dans ces conditions, le produit  $(P_i(\theta) - P_i(\eta)) \cdot (P_j(\theta) - P_j(\eta))$  est positif ou nul.

En vertu de la même hypothèse, si  $\theta \leq \eta$ , alors  $P_i(\theta) \leq P_i(\eta)$  et  $P_j(\theta) \leq P_j(\eta)$ . Ici à nouveau le produit  $(P_i(\theta) - P_i(\eta)) \cdot (P_j(\theta) - P_j(\eta))$  est positif ou nul.

En tout point du plan  $(\theta, \eta)$  le produit  $(P_i(\theta) - P_i(\eta)) \cdot (P_j(\theta) - P_j(\eta))$  est donc positif ou nul, ce qui entraîne que la covariance de  $X_i$  et  $X_j$  l'est également. Ceci achève la preuve. ■

Nous sommes maintenant en mesure d'énoncer et de démontrer des propriétés qui vont nous permettre de tester les hypothèses d'un modèle de la réponse à l'item non-paramétrique.

### PROPRIÉTÉ 3

*Si les hypothèses d'unidimensionnalité, d'indépendance locale et de monotonie sont satisfaites, alors la valeur de  $H_{ij}$  est comprise entre 0 et 1.*

**Démonstration.** Supposons que l'item  $i$  soit plus difficile que l'item  $j$ . Le coefficient d'alignement entre  $i$  et  $j$  vaut donc :

$$H_{ij} = \frac{P_{ij} - P_i P_j}{P_i(1 - P_j)}. \quad (25)$$

Le dénominateur  $P_i(1 - P_j)$  de  $H_{ij}$  est toujours positif ( $P_i \neq 0$  et  $P_j \neq 1$ ). Nous venons de démontrer que si les hypothèses d'unidimensionnalité, d'indépendance locale et de monotonie sont satisfaites, alors la covariance entre  $X_i$  et  $X_j$  est positive ou nulle. Il en découle que le numérateur  $\text{cov}(X_i, X_j) =$

$P_{ij} - P_i P_j$  de  $H_{ij}$  est positif ou nul. Ce qui entraîne que, si les hypothèses d'unidimensionnalité, d'indépendance locale et de monotonie sont satisfaites, le coefficient d'alignement  $H_{ij}$  est lui aussi positif ou nul.

Comme :

$$H_{ij} = \frac{\text{cov}(X_i, X_j)}{\text{cov}_{\max}(X_i, X_j)} \quad \text{et} \quad \text{cov}(X_i, X_j) \leq \text{cov}_{\max}(X_i, X_j), \quad (26)$$

$H_{ij}$  est toujours inférieur ou égal à 1. Il s'ensuit que  $0 \leq H_{ij} \leq 1$ . ■

#### PROPRIÉTÉ 4

*Si les hypothèses d'unidimensionnalité, d'indépendance locale et de monotonie sont satisfaites, alors la valeur de  $H_i$  est comprise entre 0 et 1.*

**Démonstration.** Supposons que les hypothèses d'unidimensionnalité, d'indépendance locale et de monotonie soient satisfaites. En vertu de la propriété précédente (Propriété 3), il s'ensuit que :

$$(\forall i)(\forall k) \quad 0 \leq H_{ik} \leq 1. \quad (27)$$

$H_i$  est plus grand ou égal au plus petit des  $H_{ik}$  (Propriété 1). Le plus petit des  $H_{ik}$  est lui-même plus grand ou égal à 0. On en tire que  $H_i$  est plus grand ou égal à 0.

$H_i$  est plus petit ou égal au plus grand des  $H_{ik}$  (Propriété 1). Le plus grand des  $H_{ik}$  est lui-même plus petit ou égal à 1. On en tire que  $H_i$  est plus petit ou égal à 1. Donc  $0 \leq H_i \leq 1$ . ■

#### PROPRIÉTÉ 5

*Si les hypothèses d'unidimensionnalité, d'indépendance locale et de monotonie sont satisfaites, alors la valeur du coefficient d'alignement global est comprise entre 0 et 1.*

**Démonstration.** Supposons que les hypothèses d'unidimensionnalité, d'indépendance locale et de monotonie soient satisfaites. En vertu de la propriété précédente (Propriété 4), il s'ensuit que :

$$(\forall i) \quad 0 \leq H_i \leq 1. \quad (28)$$

Comme  $\min_i H_i \leq H \leq \max_i H_i$  (Propriété 2), on déduit que  $0 \leq \min_i H_i \leq H \leq \max_i H_i \leq 1$ . Finalement  $0 \leq H \leq 1$ . ■

Pour que les hypothèses d'unidimensionnalité, d'indépendance locale et de monotonie soient satisfaites, il est nécessaire que les coefficients d'alignement soient positifs ou nuls. Cette condition n'est toutefois pas suffisante.

L'estimation du coefficient d'alignement global  $\hat{H}$  peut être négative pour deux raisons. La première raison est une violation de l'une ou l'autre des hypothèses du modèle de la réponse à l'item non-paramétrique. La seconde est une malencontreuse fluctuation d'échantillonnage. Lorsque l'échantillon utilisé est de petite taille, il est possible d'observer un coefficient d'alignement global  $\hat{H}$  négatif même si, dans la population, le coefficient d'alignement  $H$  est positif.

Pour trancher entre les deux termes de cette alternative, Mokken [7, pp. 160–164] propose un test statistique dont les hypothèses sont les suivantes :

$$\begin{cases} H_0 & : & H = 0 \\ H_1 & : & H < 0 \end{cases} \quad (29)$$

À partir des réponses fournies par  $n$  sujets soumis à un test contenant  $p$  items, la variable de décision  $\Delta^*$  se calcule ainsi :

$$\Delta^* = \frac{\sqrt{n-1} \cdot \Delta}{\sigma_{(p)}} \quad (30)$$

avec :

$$\Delta = \sum_{i=1}^{p-1} \sum_{j=i+1}^p (f_{ij} - f_i f_j) \quad \text{et} \quad (31)$$

$$\sigma_{(p)}^2 = \sum_{i=1}^{p-1} \sum_{j=i+1}^p f_i (1 - f_i) f_j (1 - f_j). \quad (32)$$

L'hypothèse statistique nulle  $H_0$  est rejetée au seuil  $\alpha$  si la valeur empirique de la variable de décision  $\Delta^*$  est inférieure à  $u_\alpha$ , le quantile d'ordre  $\alpha$  de la distribution normale centrée et réduite.

Pour qu'un test puisse être considéré comme une véritable échelle, il n'est pas suffisant que tous les items soient positivement corrélés et que le coefficient d'alignement global soit positif. Certains auteurs [7, 10, 11] recommandent que le coefficient d'alignement global dépasse 0.3.

Dans la section suivante, nous allons montrer, à l'aide de quelques simulations, comment se comporte le coefficient d'alignement global  $H$  dans des situations où l'hypothèse d'unidimensionnalité n'est pas toujours satisfaite.

## 2 Sensibilité du coefficient d’alignement global

### 2.1 Données

Les simulations que nous allons faire vont porter sur des matrices de données de dimension  $250 \times 12$ . Ces données imiteront les réponses de 250 sujets soumis à un test contenant 12 items dichotomiques.

Les tests que nous concocterons ne seront pas toujours homogènes. Certains contiendront 1 item intrus, d’autres en contiendront 2, 3, 4, 5 ou même 6 (*i.e.* la moitié).

Les sujets seront ainsi caractérisés selon deux traits latents  $\theta_1$  et  $\theta_2$ .

Nous étudierons l’impact de l’angle  $\gamma$  qui sépare les traits  $\theta_1$  et  $\theta_2$  en faisant varier cet angle entre  $0^\circ$  et  $90^\circ$  par saut de  $15^\circ$ . Rappelons que l’angle  $\gamma$  définit la corrélation entre  $\theta_1$  et  $\theta_2$  :

$$r(\theta_1, \theta_2) = \cos \gamma. \quad (33)$$

Le long de chaque trait, les compétences des sujets se distribueront selon une loi normale centrée et réduite :

$$\theta_1 \sim \mathcal{N}(0, 1) \quad \text{et} \quad \theta_2 \sim \mathcal{N}(0, 1). \quad (34)$$

Quant aux items, ils n’auront pas tous la même difficulté. Nous les tirerons aléatoirement d’une distribution uniforme sur l’intervalle  $[-0.75; 0.75]$ .

La probabilité qu’un sujet ayant les capacités  $\theta_j$  réponde correctement à un item  $i$  de difficulté  $\delta_i$  sera définie par une fonction logistique :

$$P_i(\theta_j) = P(X_i = 1 \mid \theta_j) = \frac{\exp(\theta_j - \delta_i)}{1 + \exp(\theta_j - \delta_i)} \quad (35)$$

avec  $j$  égal à 1 ou à 2 selon l’appartenance de l’item  $i$  aux items qui mesurent  $\theta_1$  ou  $\theta_2$ .

Notre plan de simulation se compose donc de  $6 \times 7 = 42$  cellules. Dans chaque cellule nous générerons 100 jeux de données (voir Tableau 3). Nous effectuerons ainsi un total de 4200 expériences indépendantes.

### 2.2 Traitements

Pour chaque expérience, nous calculerons le coefficient d’alignement global  $\hat{H}$  (Équation 9) et nous effectuerons un test statistique afin de savoir si le coefficient  $H$  est significativement supérieur à 0. Le test que nous allons appliquer est la version unilatérale à droite de celui que nous avons présenté à la page 10. Ainsi nous déterminerons pour chaque expérience la valeur empirique  $\delta_{emp}^*$  que prend la variable de décision  $\Delta^*$ . Rappelons que cette variable  $\Delta^*$  se distribue comme une loi normale centrée et réduite lorsque, au sein de la population, le coefficient d’alignement global est égal à 0.

Nombre d'intrus	Angle $\gamma$ entre $\theta_1$ et $\theta_2$						
	0°	15°	30°	45°	60°	75°	90°
1	100	100	100	100	100	100	100
2	100	100	100	100	100	100	100
3	100	100	100	100	100	100	100
4	100	100	100	100	100	100	100
5	100	100	100	100	100	100	100
6	100	100	100	100	100	100	100

TABLEAU 3 – Plan des simulations. Dans chaque cellule est indiqué le nombre d'expériences indépendantes réalisées.

### 2.3 Résultats

Caractérisons sommairement la distribution des coefficients d'alignement  $\hat{H}$  (voir Tableau 4).

min	$Q_1$	$Q_2$	$\bar{x}$	$Q_3$	max
0.052	0.150	0.180	0.175	0.204	0.294

TABLEAU 4 – Résumé numérique de la distribution des coefficients d'alignement  $\hat{H}$ . Les indicateurs utilisés sont par le minimum (min), le premier quartile ( $Q_1$ ), la médiane ( $Q_2$ ), la moyenne ( $\bar{x}$ ), le troisième quartile ( $Q_3$ ) et le maximum (max).

Caractérisons de la même manière la distribution des valeurs empiriques  $\delta_{emp}^*$  de la variable  $\Delta^*$  associées aux  $\hat{H}$  (voir Tableau 5).

min	$Q_1$	$Q_2$	$\bar{x}$	$Q_3$	max
7.69	22.05	26.63	25.79	30.04	41.51

TABLEAU 5 – Résumé numérique de la distribution des valeurs empiriques  $\delta_{emp}^*$  de la variable  $\Delta^*$  associées aux  $\hat{H}$ .

Si l'on décide de considérer que les items d'un test sont alignés lorsque le coefficient  $H$  est significativement positif au seuil de  $\alpha = 0.05$ , alors nous devons conclure que, dans toutes les situations que nous venons d'examiner, les items des tests sont alignés. En effet la plus petite valeur empirique  $\delta_{emp}^*$  que nous avons observée est supérieure à la valeur critique du test  $\delta_{crit}^* = u_{1-\alpha} = u_{0.95} = 1.645$ . Par construction, nous savons que cette conclusion est erronée.

Si l'on décide de considérer que les items d'un test sont alignés lorsque le coefficient d'alignement global est plus grand que 0.3 [7, 10, 11], alors nous devons conclure qu'aucun des tests que nous avons construit n'est unidimensionnel car la plus grande valeur  $\hat{H}$  observée est plus petite que 0.3. Cette

conclusion est tout aussi fausse que la précédente.

Ainsi aucun des critères classiques à notre disposition n'est suffisamment sensible pour nous permettre de distinguer les situations simulées et d'identifier les tests unidimensionnels.

Dans la partie suivante, nous allons décrire une méthode qui nous permette de regrouper les items d'un test en fonction de leur appartenance à une même sous-échelle. Du même coup, cette méthode nous permettra de savoir si tous les items d'un test sont alignés puisque, dans cette situation, tous les items du test devraient appartenir à la même sous-échelle.

### 3 Formation de bottes d'items alignés

L'algorithme que nous allons décrire s'appuie sur le fait que, dans une échelle, tous les items sont corrélés positivement (Propriété 3) d'une part et que, d'autre part, tous les coefficients d'alignement  $H_i$  sont positifs (Propriété 4).

#### 3.1 Algorithme

##### Première échelle

**Étape 1** Initialement aucun item n'a encore été sélectionné. Il s'agit ici d'identifier parmi les  $p$  items du test la paire d'items la mieux alignée, celle donc qui possède le coefficient  $H_{j_1 j_2}$  le plus grand. Ce coefficient doit être supérieur à 0 sinon la procédure s'arrête. Si plusieurs paires d'items satisfont la condition de maximalité l'une d'entre elle est choisie au hasard pour former le noyau de la première échelle.

**Étape 2** Le noyau de la première échelle est formé des items  $j_1$  et  $j_2$ . Il faut maintenant identifier parmi les  $p-2$  items restant l'item  $j_3$  qui satisfait simultanément les trois conditions suivantes :

1. l'item  $j_3$  est corrélé positivement avec l'item  $j_1$  et avec l'item  $j_2$ ;
2. le coefficient d'alignement entre l'item  $j_3$  et les deux items noyau  $(j_1, j_2)$  est positif;
3. le coefficient global  $H$  mesurant l'alignement entre les trois items  $j_1, j_2$  et  $j_3$  est maximal.

Si plusieurs items satisfont ces trois conditions l'un d'entre eux est choisi au hasard pour venir grossir le noyau de l'échelle.

**Étape 3** À cette étape, il faut identifier parmi les  $p-3$  items restant l'item  $j_4$  qui satisfait les conditions suivantes :

1. l'item  $j_4$  est corrélé positivement avec les items  $j_1, j_2$  et  $j_3$ ;
2. le coefficient d'alignement entre l'item  $j_4$  et la sous-échelle formée des items  $j_1, j_2$  et  $j_3$  est positif;
3. le coefficient global  $H$  mesurant l'alignement entre les items  $j_1, j_2, j_3$  et  $j_4$  est maximal.

Comme précédemment, si plusieurs items satisfont ces conditions, l'un d'entre eux est choisi au hasard pour venir compléter les items  $j_1, j_2$  et  $j_3$ .

**Étapes suivantes** Cette procédure se poursuit aussi longtemps que des items satisfont les conditions définies aux Étapes 2 et 3.

### Autres échelles

Au terme des étapes précédentes, s'il reste plus d'un item non sélectionné, l'algorithme reprend à l'Étape 1 afin de créer une deuxième sous-échelle. Et ainsi de suite jusqu'à ce qu'il ne soit plus possible de former de nouvelles sous-échelles. Notons que chaque item ne peut appartenir qu'à une seule sous-échelle (Voir Figure 1).

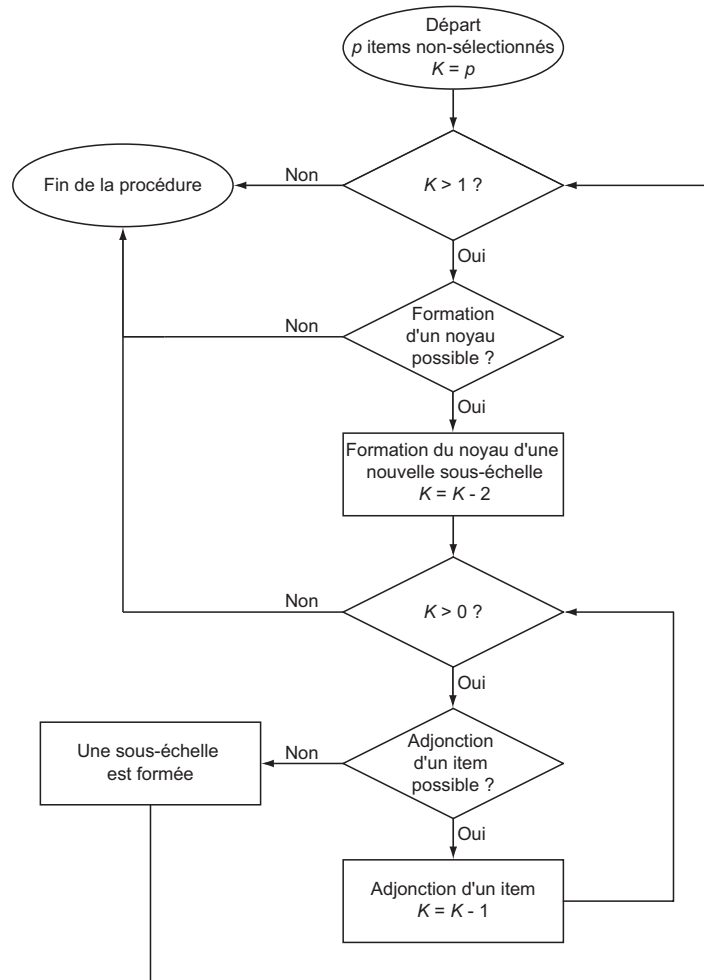


FIGURE 1 – Schéma de l'algorithme utilisé pour former des bottes d'items alignés.

### 3.2 Validation

Afin de montrer l'utilité de l'algorithme que nous venons de détailler, nous allons l'appliquer à chacun des 4200 jeux de données que nous avons générés pour éprouver la sensibilité du coefficient d'alignement global (§ 2).



Rappelons qu'il nous avait alors été impossible de distinguer les tests unidimensionnels de ceux qui ne l'étaient pas.

L'algorithme que nous venons de décrire nous permet de former des bottes d'items alignés. Au terme de chacune de ses applications nous noterons la taille de la botte la plus grosse, cette information devrait nous permettre de savoir si d'éventuels intrus ont été repérés.

Les résultats des analyses sont résumés dans la Figure 2.

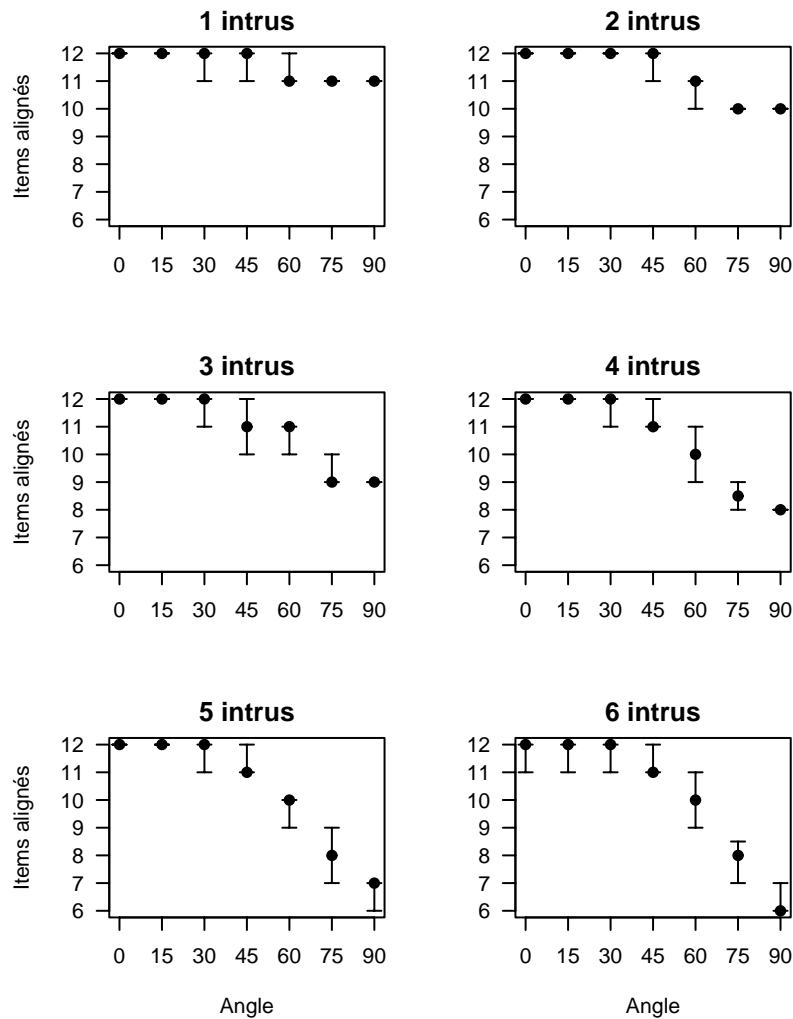


FIGURE 2 – Taille de la botte la plus grosse selon le nombre d'items intrus dans le test et l'angle formé, dans l'espace des variables, par les deux traits latents  $\theta_1$  et  $\theta_2$ . Les résultats obtenus dans chacune des 42 situations du plan de simulation sont résumés par un intervalle dont les bornes représentent respectivement le premier et le troisième quartile et un point qui indique la position de la médiane.

Lorsque l'angle, dans le plan des variables, entre les deux traits latents  $\theta_1$  et  $\theta_2$  est inférieur ou égal à  $30^\circ$ , notre algorithme ne nous permet pas de repérer les items intrus. Par contre lorsque  $\theta_1$  et  $\theta_2$  sont orthogonaux, les intrus sont quasiment toujours tous correctement identifiés et ceci quelque soit leur nombre. Entre ces deux situations quelques items intrus sont repérés et rejetés mais rarement tous.

Proposons, à titre d'essai, la règle de décision suivante : si tous les items d'un test ne se retrouvent pas dans la même botte, alors ce test n'est pas unidimensionnel<sup>5</sup>. À partir des résultats de nos simulations, calculons le nombre de fois où l'on affirmerait à tort que les items du test sont tous alignés (voir Tableau 6).

Nombre d'intrus	Angle					
	15°	30°	45°	60°	75°	90°
1	83	67	55	35	<b>9</b>	<b>1</b>
2	77	76	52	23	<b>4</b>	<b>1</b>
3	85	67	42	<b>13</b>	<b>2</b>	<b>0</b>
4	84	63	46	<b>10</b>	<b>0</b>	<b>0</b>
5	83	71	38	<b>7</b>	<b>0</b>	<b>0</b>
6	72	65	37	<b>7</b>	<b>0</b>	<b>0</b>

TABLEAU 6 – Nombre d'erreurs de seconde espèce selon le nombre d'intrus et l'angle entre les deux traits latents  $\theta_1$  et  $\theta_2$ . Dans chaque situation, 100 expériences ont été réalisées. Les taux d'erreur inférieurs à 20% sont représentés en gras.

Nous constatons que le taux d'erreur n'est inférieur à 20% que si l'angle entre les deux traits latents est supérieur à  $45^\circ$ .

Globalement, ces résultats montrent que nous sommes en mesure, grâce à l'algorithme que nous venons de décrire, d'analyser relativement finement l'alignement des items d'un test.

### 3.3 Paramètres influençant la formation des bottes d'items

Nous venons d'examiner de nombreuses situations dans lesquelles tous les items d'un même test n'étaient pas parfaitement alignés. Voyons maintenant comment certains paramètres influencent la formation de bottes d'items alignés lorsque les tests employés sont unidimensionnels.

Toutes les analyses que nous avons faites jusqu'ici ont porté sur des jeux de données de dimension  $250 \times 12$  qui correspondent à des échantillons de taille 250 caractérisés selon 12 variables. Examinons maintenant comment la taille de l'échantillon ainsi que la longueur du test influencent la grosseur des bottes d'items formées à l'aide de l'algorithme décrit au paragraphe 3.1.

<sup>5</sup>. Nous admettons que les hypothèses de indépendance locale et de monotonie sont vérifiées.

Nous générerons des échantillons de taille 125, 250 et 500. Les tests que nous créerons contiendront 6, 12, 18 ou 24 items. Nos sujets seront tirés aléatoirement d'une population dans laquelle les capacités se distribuent selon une loi normale centrée et réduite. Les items, quant à eux, seront tirés de 4 populations différentes. Dans la première population d'items, la difficulté des items se distribue uniformément entre  $-0.5$  et  $0.5$ ; dans la deuxième, la difficulté des items se distribue uniformément entre  $-0.75$  et  $0.75$ ; dans la troisième, elle se distribue entre  $-1$  et  $1$  et dans la quatrième entre  $-2$  et  $2$ .

Nous allons donc comparer  $3 \times 4 \times 4 = 48$  situations différentes. Dans chacune de ces situations, nous réaliserons 100 expériences indépendantes. Comme précédemment nous supposerons que la probabilité qu'un sujet ayant les capacités  $\theta$  résolve correctement l'item de difficulté  $\delta_i$  est égale à :

$$P(X_i = 1 | \theta) = P_i(\theta) = \frac{\exp(\theta - \delta_i)}{1 + \exp(\theta - \delta_i)} \quad (36)$$

À partir de chaque jeu de données généré, formons des bottes d'items alignés. Comptons alors le nombre de fois où tous les items se retrouvent dans la même botte (voir Tableau 7).

$p = 6$				$p = 12$			
$L$	$n = 125$	$n = 250$	$n = 500$	$L$	$n = 125$	$n = 250$	$n = 500$
1	78	<b>97</b>	<b>100</b>	1	24	<b>82</b>	<b>100</b>
1.5	68	<b>97</b>	<b>99</b>	1.5	26	<b>80</b>	<b>100</b>
2	63	<b>95</b>	<b>100</b>	2	27	71	<b>99</b>
4	56	<b>82</b>	<b>100</b>	4	10	45	<b>91</b>
$p = 18$				$p = 24$			
$L$	$n = 125$	$n = 250$	$n = 500$	$L$	$n = 125$	$n = 250$	$n = 500$
1	9	60	<b>100</b>	1	2	56	<b>99</b>
1.5	6	68	<b>98</b>	1.5	4	46	<b>97</b>
2	7	52	<b>98</b>	2	1	36	<b>94</b>
4	0	23	<b>91</b>	4	0	8	74

TABLEAU 7 – Nombre de fois où tous les items d'un test se retrouvent dans la même botte selon la taille de l'échantillon  $n$ , la longueur du test  $p$  et l'empan des difficultés des items  $L$ . Dans chaque situation 100 expériences ont été réalisées. Nous avons représenté en gras les situations dans lesquelles plus de 80% des analyses conduisent à la conclusion que tous les items sont alignés.

Pour que tous les items d'un test unidimensionnel se retrouve après l'application de notre algorithme dans une même botte, il faut que l'échantillon soit d'une taille supérieure à 250. Les résultats de nos simulations montrent qu'il est absolument inutile de réaliser ce genre d'étude avec des échantillons plus petit que 250. Lorsque l'échantillon est *grosso modo* formé de 250 sujets, les jugements d'unidimensionnalité peuvent être valables pour autant que la

gamme des difficultés des items ne soit pas trop large et coïncide avec les capacités des élèves.

La longueur du test a aussi un impact sur la formation des bottes. Toute chose égale par ailleurs, plus le test est long, moins il y a de chance que tous les items du test se retrouvent dans la même botte.

Si nous voulions être audacieux et proposer une règle garde-fou qui permette de savoir si, dans une situation donnée, le jugement d'unidimensionnalité émis suite à la formation des bottes d'items alignés est fiable, nous dirions que deux conditions sont requises : la première impose que la difficulté des items du test corresponde aux capacités des sujets et la seconde que la taille de l'échantillon soit au moins égale à 20 fois le nombre d'items du test.

## Conclusion

Nous venons de montrer qu'un emploi judicieux des coefficients d'alignement de Mokken permet d'éprouver l'unidimensionnalité d'un test complet. Pour s'assurer qu'un test incomplet [1] est unidimensionnel quelques aménagements suffisent.

Supposons que l'on ait réparti l'ensemble des items d'un test en quatre blocs disjoints et que l'on ait demandé à un premier groupe de sujets de répondre aux items des blocs 1 et 2, à un deuxième groupe de répondre aux items des blocs 2 et 3, à un troisième groupe de répondre aux items des blocs 3 et 4 et à un dernier groupe de répondre aux items des blocs 4 et 1 (voir Figure 3).

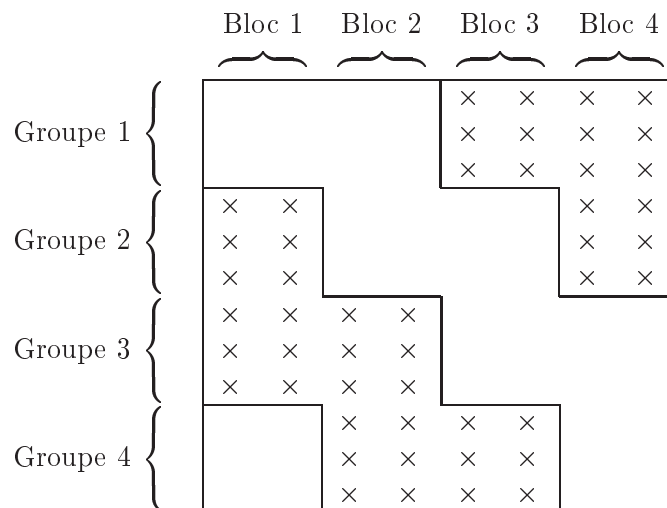


FIGURE 3 – Schéma d'un design de testage incomplet connexe. Les croix  $\times$  représentent des données manquantes par design.

Pour savoir si tous les items du test sont alignés, il faut commencer par vérifier l'alignement des items dans chaque bloc. Ensuite il faut examiner leur alignement au sein de chaque groupe, dans les paires de blocs connexes (blocs 1 et 2, 2 et 3, 3 et 4, 4 et 1). Si, lors de chacune de ces examens, l'on constate que les items sont alignés, nous pouvons alors inférer par transitivité que tous les items sont alignés et les fondements d'une application ultérieure d'un modèle de la réponse à l'item non-paramétrique aux données observées sont alors posés.

## Bibliographie

- [1] J.-P. ANTONIETTI, *Designs de testage incomplets et modèle non-paramétrique de la réponse à l'item*, Cahiers de l'IMA, 33 (2003).
- [2] D. V. BUDESCU, Y. COHEN, & A. BEN-SIMON, *A revised modified parallel analysis for the construction of unidimensional item pools*, Applied Psychological Measurement, 21 (1997), pp. 233–252.
- [3] F. DRASGOW & R. I. LISSAK, *Modified parallel analysis: A procedure for examining the latent dimensionality of dichotomously scored item responses*, Journal of Applied Psychology, 68 (1983), pp. 363–373.
- [4] R. K. HAMBLETON, H. SWAMINATHAN, & H. J. ROGERS, *Fundamentals of item response theory*, Measurement methods for the social sciences, Sage, Newbury Park, London, New Delhi, 1991.
- [5] L. G. HUMPHREYS & D. R. ILGEN, *Note on a criterion for the number of common factors*, Educational and Psychological Measurement, 29 (1969), pp. 571–578.
- [6] R. P. McDONALD, *The dimensionality of tests and items*, British Journal of Mathematical and Statistical Psychology, 34 (1981), pp. 100–117.
- [7] R. J. MOKKEN, *A theory and procedure of scale analysis with applications in political research*, De Gruyter, Mouton, The Hague, Paris, 1971.
- [8] ———, *Nonparametric models for dichotomous responses*, in Handbook of modern item response theory, W. J. Van der Linden & R. K. Hambleton, eds., Springer, New York, 1997, pp. 351–367.
- [9] R. J. MOKKEN & C. LEWIS, *A nonparametric approach to the analysis of dichotomous item responses*, Applied Psychological Measurement, 6 (1982), pp. 417–430.
- [10] K. SIJTSMA, *Methodology review: Nonparametric IRT approaches to the analysis of dichotomous item scores*, Applied Psychological Measurement, 22 (1998), pp. 3–31.
- [11] K. SIJTSMA & I. W. MOLENAAR, *Introduction to nonparametric item response theory*, Sage, Thousand Oaks, London, New Delhi, 2002.

## Table des matières

Introduction	1
<b>1 Coefficients d'alignement</b>	<b>3</b>
1.1 Définitions . . . . .	3
1.2 Propriétés . . . . .	5
<b>2 Sensibilité du coefficient d'alignement global</b>	<b>11</b>
2.1 Données . . . . .	11
2.2 Traitements . . . . .	11
2.3 Résultats . . . . .	12
<b>3 Formation de bottes d'items alignés</b>	<b>14</b>
3.1 Algorithme . . . . .	14
3.2 Validation . . . . .	15
3.3 Paramètres influençant la formation des bottes d'items . . . . .	17
<b>Conclusion</b>	<b>20</b>
<b>Bibliographie</b>	<b>21</b>