

Partie 5

La fidélité

1. Définition

- Un instrument de mesure doit permettre de différencier les sujets : les scores doivent avoir une certaine **variance**
- L'étude de la fidélité des scores est l'étude de leur variance.
Est-ce que les différences dans les scores observés sont :
 - ✓ **stables** ? Elles reflètent des différences psychologiques réelles
 - ✓ **aléatoires** ? Elles reflètent des fluctuations aléatoires
- Fidélité des scores : précision, degré d'erreur de la mesure

1. Définition

- Le modèle principal de la fidélité est la **théorie du score vrai**
 - ✓ proposée par Gulliksen (1950)
 - ✓ également appelée **théorie classique des tests**
 - ✓ cette théorie offre une définition théorique de la fidélité

2. Le modèle

- Pour une mesure donnée :
 - ✓ chaque individu se caractérise par un score vrai
 - ✓ score observé d'un sujet : son score vrai + une erreur de mesure

- Formule :

$$\mathbf{X = V + E}$$

X : score observé

V : score vrai

E : erreur de mesure

2. Le modèle

La notion d'erreur de mesure

- Postulat : l'**erreur de mesure** suit une distribution normale de moyenne égale à 0. Elle influence la variance des scores observés mais pas leur moyenne.
- L'erreur de mesure est la résultante de facteurs qui **affectent les scores de façon aléatoire**
 - ✓ facteurs individuels (humeur, état de forme, ...)
 - ✓ facteurs situationnels (environnement)
- Erreur de mesure : **bruit**

2. Le modèle

- Il existe un autre type d'erreur de mesure : l'**erreur de mesure systématique**.

On la qualifie également de **biais**

- Cette erreur de mesure est la résultante d'un (ou plusieurs) facteur(s) qui **affectent tous les scores de la même façon**
- L'erreur de mesure systématique influence la moyenne des scores observés mais pas leur variance

2. Le modèle

La notion de score vrai

- On administre 10 fois le même instrument à sujet.
A cause de l'erreur de mesure, il est très probable qu'on obtienne 10 scores différents. Ces scores suivent une distribution normale.
Quel est le « bon » score ?
- Le **score vrai** d'un sujet est la **moyenne de ses scores** (espérance mathématique) s'il passait l'instrument un grand nombre de fois
- On ne connaît donc pas le score vrai d'un sujet à un instrument

2. Le modèle

Les caractéristiques formelles

- Quatre propriétés du modèle du score vrai :

Propriété n°1

La corrélation entre les scores vrais et l'erreur de mesure est nulle

Propriété n°2

Les corrélations entre différentes erreurs de mesure sont nulles

Propriété n°3

Deux formes d'un test sont dites **parallèles** si elles ont :

- ✓ mêmes scores vrais
- ✓ mêmes erreurs de mesure

(donc mêmes moyennes et mêmes variances)

2. Le modèle

Propriété n°4

La variance des scores observés est égale à la variance des scores vrais plus la variance d'erreur

$$V_X = V_V + V_E$$

- Les différences dans les scores observés reflètent :
 - ✓ des **différences psychologiques réelles** : **variance vraie**
 - ✓ des **différences aléatoires** : **variance d'erreur**

2. Le modèle

La définition théorique de la fidélité

- La propriété n°4 permet de donner la **définition théorique de la fidélité** :

La fidélité est la proportion de variance vraie dans la variance des scores observés

$$\text{Fidélité} = V_V / V_X$$

- Cette définition est *théorique*, elle n'est pas *opérationnelle*.
(scores vrais inconnus : on ne peut pas calculer la variance vraie).
La fidélité d'une mesure ne peut pas être **calculée** mais **estimée**

3. Les méthodes d'estimation de la fidélité

- On répertorie quatre méthodes pour estimer la fidélité d'une mesure qui se différencient suivant deux critères :
 - ✓ le **nombre de formes** de l'instrument
 - ✓ le **nombre d'administrations** de l'instrument

Nombre d'administrations	Nombre de formes	
	<i>Une</i>	<i>Deux</i>
<i>Une</i>	Consistance interne	Formes parallèles sans délai
<i>Deux</i>	Stabilité temporelle (test-retest)	Formes parallèles avec délai

3. Les méthodes d'estimation de la fidélité

*La méthode de la **stabilité temporelle (test-retest)***

- Cette méthode consiste à calculer la corrélation entre les scores à un **même instrument à deux occasions différentes (t1 et t2)**
 - ✓ t1 : **test**
 - ✓ t2 : **retest**

3. Les méthodes d'estimation de la fidélité

*La méthode de la **consistance interne***

- Cette méthode générale consiste à estimer la fidélité avec **une seule forme** et **une seule administration** d'un instrument (cas le plus fréquent donc méthode la plus utilisée)
- Logique de la méthode :
 - ✓ on examine si les items de l'instrument mesurent la même chose
 - ✓ cette propriété s'appelle la **consistance interne** (ou cohérence ou **homogénéité**)

3. Les méthodes d'estimation de la fidélité

- Procédure :
 - ✓ on divise l'instrument en **plusieurs parties**
 - ✓ on considère les différentes parties comme des formes parallèles
 - ✓ on calcule les corrélations entre les parties
- Deux techniques :
 - ✓ le **split-half**
 - ✓ l'**alpha de Cronbach**

3. Les méthodes d'estimation de la fidélité

Le split-half

- Procédure :
 - ✓ on divise (**split**) l'instrument en deux parties égales (**half**)
 - ✓ on calcule la corrélation entre les deux parties
- Deux critères classiques pour créer les deux parties :
 - ✓ **Partie 1** : 1^{ère} moitié des items – **Partie 2** : 2^{nde} moitié des items
 - ✓ **Partie 1** : items pairs – **Partie 2** : items impairs
(on prend un item sur deux suivant l'ordre de passation)

3. Les méthodes d'estimation de la fidélité

- La méthode split-half présente deux inconvénients majeurs :

Inconvénient 1 : la fidélité estimée dépend du critère choisi pour créer les deux moitiés

Inconvénient 2 : la fidélité estimée est en fait sous-estimée car elle ne porte que sur la moitié des items de l'instrument

La **correction de Spearman-Brown** permet d'estimer la fidélité si on prenait en compte l'ensemble des items de l'instrument

$$\text{fidélité corrigée} = \frac{2 \times r_{A,B}}{1 + r_{A,B}}$$

avec $r_{A,B}$: corrélation entre les scores aux deux moitiés

3. Les méthodes d'estimation de la fidélité

Exemple : si la corrélation entre les scores aux deux moitiés d'un instrument est de 0.80, la fidélité corrigée est :

$$\text{fidélité corrigée} = \frac{2 \times 0.80}{1 + 0.80} = 0.89$$

3. Les méthodes d'estimation de la fidélité

- La correction de Spearman-Brown est un cas particulier de la **formule de prophétie de Spearman-Brown**

Logique :

Si la fidélité est de ... avec **m items**, quelle serait la fidélité avec **n items** ? (**m** : nombre d'items actuel, **n** : nouveau nombre d'items)

$$\text{nouvelle fidélité} = \frac{\left(\frac{n}{m}\right) \times \text{ancienne fidélité}}{1 + \left[\left(\left(\frac{n}{m}\right) - 1\right) \times \text{ancienne fidélité}\right]}$$

Exemple : avec 10 items, fidélité = 0.80

fidélité avec 15 items : 0.86

fidélité avec 20 items : 0.89

3. Les méthodes d'estimation de la fidélité

Le coefficient **alpha** de Cronbach

- Procédure :
 - ✓ dans le split-half, on divise le test en **deux parties**
 - ✓ avec le alpha de Cronbach, on divise le test en **autant de parties que d'items** : chaque item est une partie
 - ✓ on calcule soit les :
 - covariances entre les items
alpha de Cronbach **non standardisé**
 - corrélations entre les items
alpha de Cronbach **standardisé**

3. Les méthodes d'estimation de la fidélité

- Le coefficient alpha de Cronbach standardisé est le plus utilisé car il n'est pas sensible aux unités de mesure des items

On analyse les corrélations entre toutes les paires d'items :
corrélations inter-items

- Formule :

$$\alpha = \frac{n \times \bar{r}}{1 + \bar{r} \times (n - 1)}$$

n : nombre d'items

\bar{r} : moyenne des corrélations entre toutes les paires d'items

3. Les méthodes d'estimation de la fidélité

- Convention pour l'interprétation du coefficient alpha :

Valeur de α	Consistance interne
$\alpha \geq 0.9$	Excellente
$0.9 > \alpha \geq 0.8$	Bonne
$0.8 > \alpha \geq 0.7$	Acceptable
$0.7 > \alpha \geq 0.6$	Discutable
$0.6 > \alpha \geq 0.5$	Faible
$0.5 > \alpha$	Non acceptable

3. Les méthodes d'estimation de la fidélité

Bilan

- Le alpha de Cronbach augmente avec le nombre d'items.
On peut **augmenter artificiellement le alpha** en ajoutant des items redondants à l'instrument.
La valeur d'un coefficient alpha ne prend véritablement sens que si les items ne sont pas redondants, ce qui requiert donc une **analyse du contenu des items**

4. L'erreur-type de mesure

- Les scores possibles d'un sujet à un instrument se distribuent normalement autour d'une valeur moyenne (score vrai)
 - ✓ on peut calculer l'**écart-type** de cette distribution
 - ✓ si on fait de même pour l'ensemble des sujets, on peut calculer la **moyenne des écarts-types** des différents sujets
 - ✓ cet écart-type moyen est appelé l'**erreur-type de mesure** (s_e)
- Formule :

$$s_e = s_x \times \sqrt{1 - \text{fidélité}}$$

s_x : écart-type de la distribution des scores

5. Intervalle de confiance d'un score

- L'erreur-type de mesure permet de calculer l'**intervalle de confiance d'un score**
- La plupart du temps un sujet passe un test **une seule fois** : son score à cette occasion n'est probablement pas son score vrai
- La confiance dans le fait que **le score observé d'un individu correspond à son score vrai** dépend de l'erreur-type de mesure. On calcule donc un intervalle de confiance du score

5. Intervalle de confiance d'un score

- **Formule :**

$$X - z \times s_e \leq \text{score vrai} \leq X + z \times s_e$$

X : score observé du sujet

s_e : erreur-type de mesure

Intervalle de confiance à 68% : $z = 1$

Intervalle de confiance à 95% : $z = 1.96$

Exemple : score du sujet = 105 et erreur-type de mesure = 8

95% de chances que le score vrai du sujet se situe entre **89** et **120**

6. La correction d'atténuation

- Lorsqu'on calcule la corrélation entre deux mesures, la valeur de la corrélation est **atténuée** par le fait qu'une partie de la variance des mesures correspond à une variance d'erreur (la fidélité des mesures n'est jamais parfaite)
- On peut calculer une **corrélacion désatténuée** en prenant en compte la fidélité des mesures
- **Formule :**

$$r'_{XY} = \frac{r_{XY}}{\sqrt{\text{fidélité X}} \times \sqrt{\text{fidélité Y}}}$$

r_{XY} : corrélation observée entre X et Y

7. L'analyse d'items

- Afin d'améliorer la fidélité d'une mesure, on peut analyser les caractéristiques des items **a posteriori** pour sélectionner les items qui contribuent le plus à la fidélité : c'est l'**analyse d'items**
- On analyse deux caractéristiques des items :
 - ✓ la **moyenne** de l'item
On parle d'**indice de difficulté** dans le cas d'items d'aptitudes (proportion de sujets qui ont réussi/échoué à l'item)
 - ✓ l'**indice de discrimination**
Corrélation entre les scores à l'item et les scores totaux : c'est la corrélation item-total (r_{it})

Logique : un item est un « portrait miniature » du test, le score à un item doit donc apporter la même information que le score total

7. L'analyse d'items

- Les **items à sélectionner** sont ceux :
 - ✓ dont les moyennes ne sont ni trop faibles ni trop élevées
 - ✓ qui sont corrélés aux autres items (r_{it})

exemple : items de Likert en 7 points

