

Université Abderrahmane Mira-Bejaia
Faculté des Sciences Humaines et Sociales
Département de Psychologie et Orthophonie



Tests psychologiques

M1 Psychologie Clinique

7^{ème} cours

Dr GUEZOU

Année universitaire 2023/2024

La standardisation des tests

II. Les trois qualités d'un outil de collecte de données (3)

Validité

Le critère de validité est sans doute le plus complexe et on peut distinguer plusieurs formes de validité, nous retiendrons simplement :

La validité selon laquelle le test mesure bien ce qu'il est censé mesurer (construct-validity) ; un coefficient de corrélation évalue la validité par comparaison avec un autre test de même nature (et qui déjà fait ses preuves dans la pratique psychologique). Dans les années qui ont suivi sa publication, le test des cubes de Khos a été fréquemment comparé au Binet-simon ou au Terman-Merrill a été comparée au D 48 (avec des coefficients variant entre 0,60 0,80).

La validité prédictive (predictive validity) ou valeur pronostique de test, un certain nombre de tests sont censée permettre de prédire la réussite (ou l'échec) du sujet dans les situations scolaires ou professionnelles ; on compare donc la performance au test avec des évaluations de réussite effectuées.

La plupart des critiques faites au sujet des outils de collecte des données en sciences humains ont trait à la validité. Et ce souci est justifié, car la validité concerne la nature même de l'objet de la recherche. Le chercheur se demande si son outil mesure le bon chose, s'il est approprié pour mesurer le comportement étudié.

Le degré de validité d'un outil de collecte de données est également représenté par un coefficient de corrélation. Ce coefficient témoigne de la relation qui existe entre le contenu de l'outil et un critère extérieur à cet outil. La validité d'un outil ne peut être aussi élevée que sa fidélité, à cause justement de la difficulté de définir des concepts de façon opérationnelle et du degré de pertinence du critère extérieur. C'est pourquoi un outil est habituellement considéré comme valide lorsque le coefficient exprimé est de 0.50 ou 0.60.

Selon les buts de recherche, le chercheur est intéressé à établir différents types de validité de son outil de collecte de données. Habituellement, le chercheur veut mesurer, chez les participants, la présence d'un comportement (action, habiletés, attitudes, etc.) ; il est alors intéressé à établir la validité de construit, la validité de contenu et la validité concomitante de son outil. Parfois, il veut en plus prédire un comportement (ex. réussite scolaire à l'université) ; il est alors intéressé à établir la validité prédictive de son outil (ex. test d'admission à l'université).

La validité de construit

Lorsque le chercheur s'interroge sur la correspondance entre le concept qu'il veut étudier et son outil de collecte de données, il s'interroge sur la validité de construit de cet outil. Après avoir établi en quoi son concept est différent d'autres qui lui ressemblent (est-ce qu'il diffère suffisamment pour avoir une identité propre ?), il s'assure que sa démarche de traduction en dimensions puis en indicateurs est juste. La validité de construit est la plus importante des validités.

Il y a quatre façons de mesurer la validité de construit ; le chercheur peut utiliser une combinaison de ces méthodes.

- Se fier à la logique des lecteurs lorsque le concept est facilement définissable en indicateurs (ex. : traduire l'intérêt pour les études par le fait de remettre un travail dans les délais accordés).
- Solliciter l'opinion de juges experts : des personnes reconnues comme compétentes dans ce domaine particulier évaluent l'outil, sans toutefois en connaître le but ; l'outil est valide si les conclusions des juges concordent quant à cet outil semble mesurer.
- Administrer aux mêmes participants un autre outil qui mesure le même concept : l'obtention de données semblables confirme la validité du premier.
- Faire appel à un de groupe critère de personnes : ces personnes sont reconnues comme possédant beaucoup ou très peu du concept étudié ;

l'outil est valide si les données obtenues confirment les caractéristiques de ces personnes.

La validité de contenu

Lorsque le chercheur s'interroge sur la pertinence des questions ou des problèmes posés, il s'interroge sur la validité de contenu de son outil de collecte de données. Il évalue le degré de correspondance entre, d'une part, le nombre et le contenu de ces questions ou problèmes et, d'autre part, les indicateurs retenus pour représenter le concept étudié.

La validité concomitante

Lorsque le chercheur s'interroge sur la correspondance entre les données obtenues par son outil et celles obtenues par un outil qui mesure le même concept, il s'interroge sur la validité concomitante de son outil. Il doit avoir administré ce deuxième outil aux mêmes participants dans un intervalle de temps assez court pour s'assurer que le comportement étudié ne s'est pas modifié.

La validité prédictive

Lorsque le chercheur s'interroge sur la capacité de son outil de prédire un comportement chez les participants de sa recherche, il s'interroge sur la validité prédictive de son outil de collecte de données. Contrairement aux autres types de validité, la validité prédictive ne renseigne pas le chercheur sur ce que l'outil mesure ; elle permet par contre à des gens en position d'autorité de prendre des décisions quant au choix de candidats (ex. : admission à l'université, embauche de personnel).

Il va de soi que le comportement futur qui sert de critère extérieur doit être pertinent par rapport à ce que l'outil mesure. Le coefficient de corrélation exprime la relation entre les données des participants à l'outil et les comportements qu'ils ont manifestés plus tard (ex. : le lien entre les résultats d'étudiants à un test d'admission à l'université et ceux qu'ils ont obtenus à la

fin du premier cycle d'études). Plus ce coefficient s'approche de 1, plus la validité est établie ; il devient alors possible de faire des prédictions.