

La fabrication de la variabilité objective en psychologie

L'exemple de la tâche « N-back »

Stéphane Vautier et Barbara Köpke

URI Octogone-Lordat EA 4156

Université de Toulouse

- Psychologie scientifique => expliquer des phénomènes empiriques => les décrire
- Variabilité intra ou inter => pré-formatage du réel (« grilles » d'observation)
- Cadres descriptifs :
 - arbitraires
 - réductifs
 - théoriquement contraignants (cf. Windelband)

C'est précisément dans la mathématique que la relation de la connaissance à ses objets, formés par cette connaissance même, possède une transparence extrêmement instructive. Car, si les configurations numériques et les figures spatiales particulières que le chercheur constitue en objets dépendent de son arbitraire, la pensée connaissante, dès que l'objet est construit, est pourtant complètement liée à lui et soumise à la légalité immanente de l'objet. La production du problème et sa solution ne peuvent rien viser d'autre que le développement des rapports de grandeur créés par la construction synthétique même. La pensée mathématique, aussi libre qu'elle soit dans la production de ses objets, fait précisément en cela l'expérience de la *contrainte de l'objectivité* [*Gegenständlichkeit*], qui se trouve contenue dans les configurations qu'elle construit et qui oppose sa puissance souveraine à toute affirmation arbitraire. Nous voulons en outre qualifier cette relation, qui apparaît nécessairement du point de vue de la psychologie subjective comme un miracle, de *logique de l'objet*.

Il en va toutefois de même pour les *sciences empiriques*. Si les habitudes de la pensée préscientifique, relevant du réalisme naïf, le dissimulent, la réflexion logique doit du coup d'autant plus rigoureusement mettre en évidence le fait que nulle connaissance ne peut se rapporter à des objets qu'elle n'a pas elle-même conceptuellement déterminés. (p. 222)

Windelband, W. (1912/2002). *Qu'est-ce que la philosophie ? et autres textes*. Paris : Vrin.

Argument épistémologique

- La contrainte logique du cadre descriptif est méconnue en psychologie ;
- sinon
 - on ne présenterait pas certains tests psychotechniques ou tâches expérimentales comme des dispositifs de **mesure** ;
 - on adopterait une autre approche de la variabilité à laquelle ces techniques descriptives donnent accès.

Application : la tâche « N-back »

- Famille de tâches expérimentales pour « mesurer la mémoire de travail (MT) »
- Exemple : Jaeggi, S. M., Buschkuohl, M., Perrig, W. J., & Meir, B. (2010). The concurrent validity of the N-back task as a **working memory measure**. *Memory*, 18, 394-412.
- Principe (N = 2)

A B A D E D A E A

A B A D E D A E A

Construction des scores

Valeur de la réponse Temps de réaction

A +1 si A t_1 si A

D -1 si lettre

E -1 si lettre

D +1 si D t_2 si D

A -1 si lettre

...

→ $\text{Score}_1 = \text{nb}(+1) - \text{nb}(-1)$: entier relatif.

→ $\text{Score}_2 = \sum(t_i) / \text{nb}(+1)$: réel positif.

A B A D E D A E A

Significations des scores

→ $\text{Score}_1 = \text{nb}(+1) - \text{nb}(-1)$: entier relatif.

→ $\text{Score}_2 = \sum(t_i) / \text{nb}(+1)$: réel positif.

- Evaluative 😊 (<, >, =)

Paul a une meilleure performance que Julie.

Paul a une meilleure performance aujourd'hui qu'il y a 15 jours.

- Descriptive 😞

On ne sait pas spécifier les observations d'origine.

Mesurer versus évaluer

- On sait mesurer une **grandeur théorique** ($q =$ quantité de capacité de mémoire de travail) grâce à des **observations** (y) s'il existe F croissante telle que $y = F(q)$ parce qu'on veut

$(y < y' \Rightarrow q < q')$ = argument valide.

- On sait évaluer une **performance** ($x =$ un score) grâce à des **observations** (y) en se donnant une fonction A (agrégation) telle que $x = A(y)$ parce qu'on veut

x et x' comparables : $x \leq x'$ ou $x \geq x'$.

- Mesurer c'est expliquer : F **existe** : testable ?
- Scorer c'est évaluer : on **institue** A : consensus ?

Problèmes

1. Faux problème : les **scores** sont-ils des mesures ? (Non ce sont des évaluations.)
2. Faux problème : comment expliquer les **scores** ? (Par la règle d'agrégation qu'on s'est donnée et les observations qu'on a effectuées.)
3. Problème mal posé : comment expliquer les relations statistiques des variables de scores avec d'autres variables ? (Quelles explications peut-on tester ?)
4. Problème : comment tirer parti, dans la pratique, des scores ? (Validité prédictive.)
5. Sait-on mesurer la **capacité de la mémoire de travail** ?

Sait-on mesurer la **capacité de la MT** ?

Oui, grâce aux **scores** ?

- Notations : (x_1, x_2) un 2-uplet de **scores** formés par split-half.
- Hypothèse : il existe $F = F_1 F_2$ croissante qui, à toute quantité q de **capacité de MT**, associe un unique (x_1, x_2) .
- Csq. : quels que soient $F(q) = (x_1, x_2)$ et $F(q') = (x'_1, x'_2)$,
 $(x_1, x_2) \leq (x'_1, x'_2)$ si $q \leq q'$
ou $(x_1, x_2) \geq (x'_1, x'_2)$ si $q \geq q'$.
- S'il existe $(x_1, x_2) <> (x'_1, x'_2)$, F n'existe pas.

Probablement « non, pas grâce aux scores » (question empirique).

Sait-on mesurer la capacité de la MT ?

Oui, grâce aux **observations** ? Qu'observe-t-on ?

Réponse

- A « **oui** » (1, correct) ou « **pas oui** » (0) et t_1
- D « **oui** » (0) ou « **pas oui** » (1, correct) et t_2
- E « **oui** » ou « **pas oui** » (correct) et t_3
- D « **oui** » (correct) ou « **pas oui** » et t_4
- A « **oui** » ou « **pas oui** » (correct) et t_5

Sait-on mesurer la capacité de la MT ?

- Hypothèse : étant donné un répondant, il existe

$$F = F_1 F_2 \dots F_m : [0, \max] \rightarrow (0, 1)^m$$

qui, à tout q , associe un unique m -uplet de réponses.

- Conséquence : soit y et y' les m -uplets d'un répondant en test-retest ; y et y' sont comparables.
- S'il existe $y \leftrightarrow y'$, F n'existe pas.

Probablement « non, pas grâce aux observations » (question empirique).

Etat des connaissances

- Les réponses à la tâche N-back sont :
 - des m -uplets de 2-uplets comme
$$((1, t_1), (0, t_2), (1, t_3), \dots, (1, t_m)),$$
 - des m -uplets comme
$$(1, 0, 1, \dots, 1),$$si on ignore les temps de réponse.
- Expliquer les réponses d'un répondant, c'est d'abord identifier les **signes** de contraintes qui s'exerceraient sur la variabilité de **ses** réponses dans $\{0, 1\}^m \times \{0, 1\}^m$
- Ce qui nécessite $n \gg 2^{2m}$ observations.

	0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111	
0000																	16
0001																	16
0010																	16
0011																	16
0100																	16
0101																	16
0110																	16
0111																	16
1000																	16
1001																	16
1010																	16
1011																	16
1100																	16
1101																	16
1110																	16
1111																	16
																	256

Les signes de contrainte sont les « trous » dans une ligne dont la marge est >> 16.

- Trouver une loi expérimentale (au moins une fréquence conditionnelle nulle) => disposer d'une information en quantité suffisante (n très grand).
- Pas assez d'information => pas de découverte de loi expérimentale.
- Faute de loi expérimentale, on peut chercher des lois de probabilité (conditionnelles).
- Estimer une loi de probabilité cond. => distribution de fréquences **robuste** => n très grand.
- Si n petit, on ne connaît pas bien le phénomène (donc prématuré de chercher à l'expliquer).
- Si n ne peut pas être grand, notre rôle de scientifiques est de vulgariser l'**ignorance objective** dans laquelle on se trouve parce que les phénomènes ne sont pas bien documentés : réalisme scientifique => respect du **principe de réalité** (tout n'est pas possible).

Retour au scorage (évaluation) : approche bivariée de la performance

Les observations d'un répondant permettent de **calculer** deux indices statistiques (de **rendement**) qui positionnent sa performance dans $[0, 1]^2$.

	Présent	Absent
Oui	a	b
Pas oui	c	d

- $Se = a / (a + c)$
fq(Oui | Présent)

- $Sp = d / (b + d)$
fq(Pas oui | Absent)