

REPRÉSENTATIONS SOCIALES DE LA STATISTIQUE CHEZ DES ÉTUDIANTS DE PSYCHOLOGIE, MISES EN ÉVIDENCE À TRAVERS LES RÉSEAUX D'ASSOCIATIONS

Jean-Marie Marion¹, Véronique Dubreil-Frémont², Alain Bihan-Poudec³ et Noëlle Zendreras⁴

¹ Université Catholique de l'Ouest – Faculté des Sciences – BP 10808 – 49008 ANGERS CEDEX 01 jean-marie.marion@uco.fr

² Université Catholique de l'Ouest – Faculté des Sciences Humaines et Sociales – BP 10808 – 49008 ANGERS CEDEX 01 veronique.dubreil@uco.fr

³ Université Catholique de l'Ouest – Faculté d'Éducation – BP 10808 – 49008 ANGERS CEDEX 01 alain.bihan-poudec@uco.fr

⁴ Université Catholique de l'Ouest – Faculté des Sciences Humaines et Sociales – BP 10808 – 49008 ANGERS CEDEX 01 noelle.zendreras@uco.fr

Résumé. En tant qu'enseignants en statistique dans l'enseignement supérieur, nous nous intéressons aux représentations sociales qu'ont nos étudiants à propos de cette discipline. Une première série d'études à base de questionnaires nous a permis de cerner les représentations sociales de la statistique auprès d'étudiants en sciences humaines et sociales (Bihan-Poudec, 2013), ainsi que leur évolution (Bihan-Poudec et Marion, 2014). Plus récemment, nous avons poursuivi nos travaux chez des étudiants en psychologie, à l'aide de la technique des réseaux d'associations (de Rosa, 1995) afin d'explorer notamment la dimension affective de l'attitude des étudiants vis-à-vis de la statistique (Dubreil-Frémont, Marion et Bihan-Poudec, 2018). Par ailleurs, comme son nom l'indique, cette technique permet aux participants d'associer eux-mêmes les mots, et aux expérimentateurs de ne plus se contenter de simples co-occurrences.

Mots-clés. Recherche en éducation statistique, représentations sociales, attitudes, réseaux d'associations

Abstract. As statistics teachers, we are concerned with undergraduate students' attitudes toward statistics. In order to measure students' attitudes toward statistics and its progression, we first used questionnaires that were administrated to undergraduate students in social sciences (Bihan-Poudec, 2013; Bihan-Poudec & Marion, 2014). Then, we used an associative network technique (de Rosa, 1995) to investigate the affect component of content of students' attitudes. As the name suggests, this technique enables people to link words they think it should be, and avoids experimenters settling for occurrences.

Keywords. Statistics education research, social representations, statistics attitudes, associative networks

1. Introduction

Dans la littérature internationale, les représentations sociales et les attitudes manifestées par les étudiants au regard de la statistique ont fait l'objet de nombreuses études dont Emmioğlu et Capa-Aydin (2012) fournissent une méta-analyse. La statistique, apprenants, professionnels et citoyens peinent à se la représenter et à la définir. Comme nous le mettons en exergue (Zendreras, Dubreil-Frémont, Marion et Bihan-Poudec, 2017), les spécialistes eux-mêmes en proposent des approches et

des définitions très variées.

2. La statistique ? Des mathématiques, des pourcentages, des chiffres, des calculs...

« Mathématiques », « pourcentages », « chiffres », « calculs »... tels sont les termes les plus fréquemment utilisés sous une forme ou une autre quand un étudiant en sciences humaines et sociales est invité à décliner ce que la statistique évoque pour lui avant même son premier cours dans cette discipline (Bihan-Poudec, 2013). Par exemple le mot « calcul » est évoqué par 20,5 % des 614 étudiants de la recherche; quant au terme « mathématiques », il est cité par plus d'un tiers des enquêtés, tout comme « pourcentage ». Nous avons donc un champ lexical partagé (*shared representation*), dans la mesure où les vingt premiers lemmes les plus utilisés représentent près de deux tiers des associations des étudiants. Mieux, si nous nous arrêtons sur les deux lemmes les plus fréquents – « mathématiques » et « pourcentage » –, 247 sujets sur 614 (40 %) ont cités soit l'un, soit l'autre, soit les deux, dans leurs deux premières associations.

3. Une unanimité mise à mal ?

Toutefois, ce constat de discours partagé mérite d'être nuancé par plusieurs observations. Il convient de souligner la production hétérogène des sujets : à un étudiant correspondait 3,82 unités lexicales en moyenne (avec un écart-type de 2,41) mais cela varie d'1 seul mot donné à 18, la moitié des étudiants donnant entre 2 et 5 mots. De fait, dès lors, il apparaît difficile d'avoir des associations originales au vu de cette production relativement restreinte. Quant aux nombres d'occurrence des unités lexicales, ils varient de 1 à 96 : si une unité est mentionnée 2,14 fois en moyenne, l'écart-type est très élevé (6,59), interrogeant la fiabilité de la moyenne comme indice de tendance centrale (le coefficient de variation s'établit à 308 %).

Enfin, il existe des sujets qui ne partagent pas le discours commun. S'il est logique que deux étudiants sur trois n'ait pas utilisé le terme « mathématiques » (puisqu'un tiers le font), il s'avère que 45,8 % n'utilisent ni « mathématiques », ni « pourcentage », termes les plus fréquents. Plus globalement, il existe une proportion d'étudiants (9,3 %) qui n'ont jamais eu recours à l'un des 20 lemmes les plus utilisés ; un quart à ne recourir à aucun des cinq premiers mots dans la totalité de leurs associations ; plus d'un tiers n'a utilisé aucun des trois premiers mots.

De l'ensemble de ces remarques en ressort, non pas la preuve, mais la possibilité que derrière cette apparence de discours partagé, existeraient des sous-discours, pourquoi pas eux aussi homogènes, et qui correspondraient à des caractéristiques particulières d'étudiants.

4. Statistique ? Parle-t-on de la même chose ?

À cette fin, nous avons caractérisé les étudiants en fonction de la formation qu'ils suivaient. Au sortir d'analyses factorielles (Bihan-Poudec, 2013), il en apparaît plusieurs types de positionnement vis-à-vis de la statistique.

- Pour les étudiants en Psychologie, l'intention d'étude est prégnante et l'ancrage dans les mathématiques fort. Ce rapport académique donc à la statistique, dérivée des *mathématiques appliquées* à des études, est renforcé par la relative absence de termes opératoires, tels RECUEIL, pas de COMPRÉHENSION ou d'INTERPRÉTATION.
- Ce sont ces termes qui prédominent pour les étudiants *adultes* en Licence 3 Sciences de l'Éducation, la *finalité d'analyse* est ainsi présente, réflexive. Se trouve évincée la dimension mathématique reliée à une démarche de recherche « universitaire ».
- Un autre positionnement apparaît plus général, d'intention : ainsi, pour les étudiants en L1 Sciences de l'Éducation, avec la statistique conçue comme *démarche d'étude* : l'on observe, l'on étudie une partie d'une population et l'on en voit les résultats.

- Mais cette intention peut même être absente, comme chez les jeunes étudiants en Licence 3 Sciences de l'Éducation où la statistique est *réifiée* et se résume laconiquement aux statistiques, entendues comme simples informations, DONNÉES et RÉSULTATS, sans guère de finalité.

5. Les réseaux d'associations chez les étudiants en psychologie

Même si nous avons montré que chez les étudiants en psychologie ces images de la statistique n'évoluaient guère après l'enseignement de cette discipline (Bihan-Poudec et Marion, 2014), il n'en demeure pas moins que ces positionnements vis-à-vis de la statistique sont des constructions faites à partir de la convocation par les analyses factorielles des occurrences des termes utilisés par les étudiants : ces derniers n'en sont aucunement directement les auteurs. C'est pour cela que nous avons eu recours à la technique des réseaux d'associations mise en œuvre par de Rosa (1995).

Pour la présente recherche, des étudiants en psychologie étaient invités à inscrire les mots qu'évoquaient pour eux le terme de « statistique », puis à relier entre eux les mots qui leur semblaient être liés. Dans certains cas, les étudiants ont même explicité ces associations.

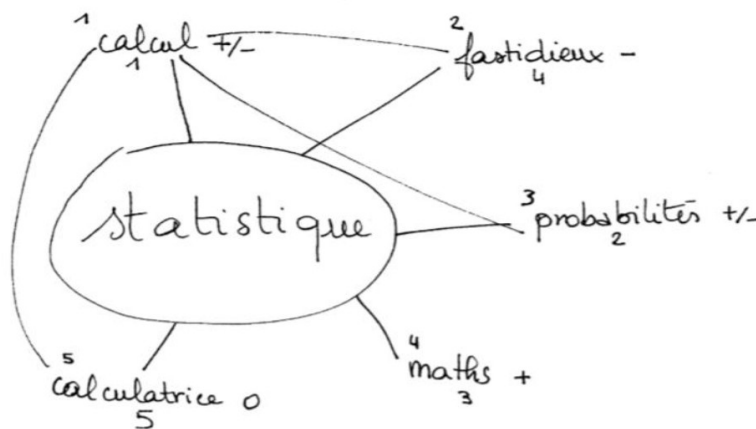


Figure 1 : Exemple de réponse d'un étudiant.

Les 142 étudiants ayant participé ont écrit entre 4 et 14 mots évoqués par « statistique », soit une moyenne de 7,7 mots avec un écart-type de 2,2. Les mots les plus fréquemment cités étaient « mathématiques », « nombres » et « calculs », comme observé précédemment (*vide supra*). Ils ont cité au total 1100 mots. Nous avons donc procédé à un regroupement de mots ayant la même signification, puis focalisé sur les 32 mots cités plus de 8 fois.

5.1 Rang d'importance et connotations des mots évoqués par le terme « statistique »

Nous avons demandé aux étudiants d'attribuer un ordre de préférence aux mots évoqués, ce qui nous a permis de déterminer le rang moyen associé à chacun d'eux. Dès lors, les mots apparaissant les plus importants sont « recherche », « méthode », « sondages-enquêtes », « population », « étude », « mathématiques »... tandis que les moins importants étaient « variance », « diagramme » et « écart-type ». Les étudiants ont aussi eu à noter chaque mot à l'aide de signes +, -, 0 ou ±, selon que le mot en question ait une connotation respectivement positive, négative, neutre ou ambiguë. Ainsi pour les 32 mots les plus cités, un indice de polarité a été calculé selon la formule

$$\text{indice de polarité} = \frac{\text{nombre de mots positifs} - \text{nombre de mots négatifs}}{\text{nombre total de mots}}$$

Les mots ayant l'indice de polarité le plus fort c'est-à-dire ceux qui sont perçus de façon positive par les étudiants de psychologie sont « méthode », « utile », « sondages-enquêtes » et « recherche », alors que les mots connotés le plus négativement sont « variance », « fastidieux » et « difficultés ».

incompréhension ».

L'ordre d'importance des mots et leur indice de polarité ont permis d'élaborer un classement des 32 mots les plus cités.

Mot	Indice associant polarité et importance du mot
méthode	3,53
utile	2,77
recherche	2,63
sondages - enquêtes	2,54
population	1,92
étude	1,61
expérience	1,10
échantillon	1,03
logique	0,94
filière	0,92
probabilité	0,86
comparaison	0,85
scientifique	0,71
cours	0,71
moyenne	0,70
tableaux	0,65
nombres	0,29
variable	0,28
médiane	0,18
graphique	0,16
fréquence	0,13
quartile	0,09
calculatrice	0,08
diagramme	0,01
donnée	0,00
écart type	0,00
variance	-0,08
formule	-0,35
calcul	-0,51
mathématiques	-0,79
fastidieux	-1,53
difficultés - incompréhension	-1,63

Tableau 1 : Classement des mots évoqués par le terme « statistique » associant importance et polarité connotative.

Ce classement montre que les mots à connotation positive et considérés comme importants sont liés aux études, enquêtes et recherches, et au caractère « utile » de la statistique en tant que « méthode ». À l'opposé, les « mathématiques » sont associées aux « calculs » et aux « difficultés » rencontrées par les étudiants de psychologie.

5.2 Liaisons entre les mots établies par les sujets

Au cours de la passation, chaque sujet a pu relier entre eux les mots qui lui semblaient être liés, ce qui a permis d'obtenir un réseau d'associations.

Les liaisons établies par l'ensemble des sujets, les plus fréquentes, associent des mots liés à la

statistique descriptive.

Mot 1	Mot 2	Nombre de liaisons
mathématiques	calcul	33
mathématiques	nombres	22
mathématiques	difficultés - incompréhension	15
moyenne	quartile	14
moyenne	médiane	14
nombres	calcul	12
mathématiques	fréquence	11
nombres	fréquence	10
calcul	fréquence	9
calcul	moyenne	9
tableaux	graphique	9
mathématiques	calculatrice	8
calcul	calculatrice	8
cours	écart-type	8

Tableau 2 : Couples de mots liés le plus fréquemment.

Sur les 142 étudiants, « mathématiques et « calcul » ont été reliés 33 fois.

En rapportant le nombre de liaisons entre deux mots chez l'ensemble des sujets avec leur évocation simultanée chez un même sujet, il a été possible de déterminer leur fréquence de liaison.

$$\text{pourcentage de liaisons} = \frac{\text{nombre de liaisons entre 2 mots}}{\text{nombre de fois où ces 2 mots sont cités simultanément}}$$

Les couples de mots qui engendrent le plus de liaisons dès lors qu'ils sont cités simultanément sont présentés dans le tableau suivant :

Mot 1	Mot 2	Fréquence de liaisons
calculatrice	moyenne	83,33%
mathématiques	calcul	71,74%
mathématiques	difficultés - incompréhension	68,19%
calcul	calculatrice	61,54%
nombres	probabilité	57,14%
graphique	tableaux	56,25%
graphique	fréquence	55,56%
nombres	fréquence	52,63%

Tableau 3 : Pourcentages de liaisons observées pour les couples de mots cités simultanément.

Ainsi le couple « mathématiques » « calcul » est cité 46 fois et relié à 33 reprises.

Ces associations spontanées réalisées par les étudiants enrichissent la description de leur représentation de la statistique. Elles viennent confirmées les difficultés ressenties par les étudiants en psychologie vis-à-vis de la statistique notamment en termes d'opérations mathématiques.

Par la suite, nous présenterons quelques autres résultats obtenus, en particulier nous commenterons une analyse d'échelles tenant compte des indices de polarité et des indices de neutralité, suivant ainsi le point de vue de de Rosa (1995), puis nous présenterons une analyse des proximités appliquée à un tableau de similarité entre les mots proposés par les étudiants pour qualifier la statistique, ceci en utilisant des méthodes de *multidimensional scaling*.

Dans cette étude, notre intention a été triple : celle d'approfondir les rapports des étudiants à la statistique ; celle de voir si leur positionnement coïncide avec celui de leurs prédécesseurs, qualifié

d'académique, de mathématique ; enfin, celle de valider (ou non), le principe de reconstruction de *discours* à partir de l'analyse statistique de *mots*.

Bibliographie

- Bihan-Poudec, A (2013). *Des chiffres et des êtres. Étude introductive à l'identification de la représentation sociale de la statistique chez des étudiants de premier cycle en Sciences humaines et sociales en France*, (thèse de doctorat en Éducation (PhD) non publiée). Université de Sherbrooke, Québec, Canada. <http://iase-web.org/documents/dissertations/14.AlainBihanPoudec.Dissertation.pdf>
- Bihan-Poudec, A. et Marion, J.-M. (2014). *De l'évolution de la représentation de la statistique chez des étudiants en Sciences humaines et sociales*, 46^{es} Journées de Statistique, Société Française de Statistique, Rennes, 5 juin.
- Dubreil-Frémond, V., Marion, J.-M. et Bihan-Poudec, A. (2018). *Attitudes d'étudiants en psychologie envers la statistique*, 50^{es} journées de Statistique, Société Française de Statistique, Saclay, 1^{er} juin.
- de Rosa, A. S. (1995). Le « réseau d'associations » comme méthode d'étude dans la recherche sur les représentations sociales : structure, contenus et polarité du champ sémantique, *Les Cahiers Internationaux de Psychologie Sociale*, 28, 96-120.
- Emmioglu, E. et Capa-Aydin, Y. (2012). Attitudes and Achievement in Statistics: A Meta-Analysis Study, *Statistics Education Research Journal*, 11(2), 95-102. https://www.stat.auckland.ac.nz/~iase/serj/SERJ11%282%29_Emmioglu.pdf
- Zendreras, N., Dubreil-Frémond, V., Marion, J.-M. et Bihan-Poudec, A. (2017). Les données et leur production: réflexions sur une lacune paradoxale en éducation statistique, *Statistique et Enseignement*, 8(2), 59-78. <http://statistique-et-enseignement.fr/article/view/644>