



Compte rendu du colloque "Méthodes Quantitatives en Sciences Humaines (MQSH) 2024"

Pier-Olivier Caron ^a , Olivier Vivier ^b , Eva Lefebvre ^b , André Achim ^b , Ibtissem Ben Alaya ^b , Eric Frenette ^c , Louis Laurencelle ^d & Sophie Vanbelle ^e

^aUniversité TÉLUQ

^bUniversité du Québec à Montréal

^cUniversité Laval

^dUniversité du Québec à Trois-Rivières

^eUniversité de Maastricht

Abstract ■ Le vendredi 7 juin 2024 s'est tenu le 13e colloque annuel Méthodes quantitatives en sciences humaines (MQSH) à l'Université TÉLUQ à Montréal. Cinq conférenciers et conférencières ont présenté leurs travaux dans le domaine. Eric Frenette a discuté de l'évolution du modèle Bifacteur, depuis sa proposition par Holzinger, suite aux travaux de Spearman en 1927, jusqu'à sa version contemporaine, version digne du monstre de Frankenstein, qu'elle est devenue en y transfigurant ses hypothèses sous-jacentes et en y rapiéçant des paramètres augmentant son potentiel explicatif. Sophie Vanbelle, conférencière invitée pour l'évènement, a exposé les principales mesures statistiques et leur interprétation pour déterminer la fiabilité et l'accord dans le cas d'une échelle binaire. Elle a discuté de l'importance de rapporter les différents indices d'accord, de calculer l'effectif d'échantillon nécessaire pour avoir des mesures fiables de l'accord interjuges (et intrajuges) et du besoin de lignes directrices pour les chercheurs et chercheuses. André Achim a proposé une approche novatrice de l'analyse factorielle exploratoire reposant non pas sur la décomposition de matrices, mais plutôt sur l'annulation du signal, ce qui permet d'extraire toutes les informations pertinentes, le nombre de facteurs et la structure factorielle, tout en réduisant le nombre d'options décisionnelles. Ibtissem Ben Alaya a présenté le processus de validation en sept étapes d'un questionnaire portant sur les compétences socioémotionnelles des enseignants et enseignantes tout en discutant des défis rencontrés ainsi que des choix psychométriques pris. Louis Laurencelle, dans sa présentation, a critiqué le raisonnement et les critères d'altération du taux d'erreur de type I (seuil α), telles la correction de Bonferroni et celle de Tukey, lesquels torpillent la puissance statistique des travaux de recherche en sciences humaines. Il a examiné et argumenté la nocivité d'une telle approche et son absurdité. En conclusion, cette 13e édition du colloque MQSH fut digne de sa bannière : compétence, critique et créativité.

Acting Editor ■ Denis Cousineau (Université d'Ottawa)

English abstract at the end.

Keywords ■ Statistiques, Actes de colloque, MQSH 2024.

pier-olivier.caron@teluq.ca

[10.20982/tqmp.20.2.p156](https://doi.org/10.20982/tqmp.20.2.p156)

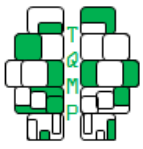
Introduction

Suite à la résurrection du colloque Méthodes Quantitatives en Sciences Humaines (MQSH) en 2023 (Caron et al., 2023), une vingtaine de participants se sont réunis le vendredi 7 juin 2024 pour la 13e édition du colloque annuel à l'Université TÉLUQ, Montréal. Le présent article fait un compte rendu des présentations et discussions ayant eu

lieu lors de l'évènement. La structure de l'article respecte le programme de la journée.

Mot de bienvenue

Pier-Olivier Caron, organisateur, débute la journée en s'adressant aux étudiants et étudiantes, professionnels et professionnelles, chercheurs et chercheuses en méthodes quantitatives. Il rappelle d'abord l'objectif du colloque, soit



de promouvoir tous ceux et toutes celles qui exercent la fonction de conseillers et conseillères en méthodes quantitatives en proposant un lieu de rencontre où s'échange l'expertise francophone en méthodes quantitatives en sciences humaines. Possédant une expertise à mi-chemin entre les mathématiques statistiques et les sciences sociales, ces chercheurs et chercheuses sont indispensables à la bonne conduite de la recherche scientifique, particulièrement avec les récentes controverses dans le monde scientifique. Or, les méthodes quantitatives sont parfois considérées comme étant accessoires, un simple outil permettant d'atteindre la publication d'un article scientifique. L'expertise en méthodes quantitatives est mal reconnue parmi les collègues, malgré la diversité, l'originalité et la pertinence des contributions qu'un conseiller ou qu'une conseillère en méthodes quantitatives peut apporter. Pier-Olivier termine en prenant un moment pour remercier l'expertise des membres présents à ce colloque, et pour souligner leur apport fondamental à la science.

Le modèle Bifacteur : de sa genèse à la version Frankenstein

Pour débiter sa conférence, Eric Frenette (Université Laval) rappelle que développer un questionnaire est un processus complexe qui repose sur l'obtention d'un maximum de preuves de validité. À cause de formalités rébarbatives, l'étude de la validité d'un questionnaire est souvent remplacée par des heuristiques statistiques comme les indicateurs de fiabilité, au profit d'une étude complète de la structure interne et des composantes de validité nécessaire à l'élaboration du questionnaire. Lors de la vérification de la structure factorielle, le modèle Bifacteur confirmatoire est de plus en plus utilisé pour appuyer l'ajustement de modèles théoriques aux données. Pour étayer ce point, Frenette présente l'évolution du nombre de publications concernant le Bifacteur, passablement ignoré pendant près de 70 ans avec moins de 10 publications le concernant par décennies (entre 1930 et 1999). À partir des années 2000, plus de 3000 publications sont référencées par décennies.

Sur cette base, Frenette questionne l'utilisation contemporaine du modèle Bifacteur et, pour ce faire, présente un retour historique. Il débute en introduisant le modèle à deux facteurs de Spearman (1927), duquel le modèle Bifacteur de Holzinger (1935) est un prolongement. La théorie sous-jacente du modèle de Spearman suggère qu'une habileté générale explique toutes les mesures (i.e., unidimensionalité) et que plusieurs habiletés spécifiques indépendantes existent (une par mesure). Le modèle à deux facteurs de Spearman utilise une approche statistique fondée sur les différences de

tétrades, un test qui vise à déterminer l'adéquation d'un facteur commun. Les différences de tétrades permettent d'identifier des regroupements entre les facteurs spécifiques qui ne sont plus indépendants, des facteurs de groupe selon Spearman. C'est sur ces principes qu'en 1935 Holzinger propose le modèle Bifacteur. Il suggère de remplacer les différences de tétrades par l'analyse de la matrice résiduelle. Ce modèle comporte des conditions d'application : l'unidimensionalité et l'analyse de la matrice résiduelle suite au retrait du facteur dominant pour identifier les facteurs de groupe.

Frenette termine en abordant l'utilisation actuelle du modèle Bifacteur confirmatoire en le comparant aux conditions d'application à l'origine du modèle Bifacteur proposé par Holzinger et en discutant des modifications insidieuses s'étant glissées dans cette analyse au fil du temps, la dénaturant au point où Frenette la qualifie de version Frankenstein. La monstruosité de l'application de cette version contemporaine provient de son inconformité aux fondements théoriques originaux. En fin de compte, le modèle Bifacteur a évolué sous une nouvelle forme qui ne respecte pas ses origines, ses caractéristiques et surtout ses conditions d'application. Ces transformations lui permettent un pouvoir d'explication dissimulé en le surchargeant par l'ajout d'un nombre substantiel de liens lui permettant d'obtenir des statistiques d'ajustement appréciables.

Mesures de fidélité et d'accord pour données binaires

La conférencière invitée, Sophie Vanbelle (Université de Maastricht), examine les principales mesures statistiques utilisées pour déterminer la fiabilité et l'accord et leur interprétation dans le cas d'une échelle binaire dans le cadre médical (voir Hernandez Engelhart et al., 2023, pour le contexte). Bien que des lignes directrices générales existent pour planifier et analyser des études de fiabilité et d'accord, les chercheurs et chercheuses sont pris au dépourvu devant la diversité des mesures statistiques, leurs différentes méthodes d'inférence statistique et les lacunes quant à la planification comme la détermination de la taille d'échantillon (nombre de sujets, de juges, de répétitions par juge). En pratique, les chercheurs et chercheuses vont principalement recourir aux mesures statistiques disponibles dans le programme informatique qu'ils utilisent, comme le κ de Fleiss dans SPSS. Ainsi, la présentation de Vanbelle vise à discuter des différentes pratiques liées à l'analyse des mesures de fidélité et d'accord ainsi que d'améliorer leur processus d'inférence statistique.¹

Pour exemplifier les mesures d'accord interjuges, Vanbelle utilise le cas le plus simple pour examiner les mesures d'accord interjuges, soit une échelle binaire à deux juges.

¹La présentation est partiellement basé sur Vanbelle (2024).



Dans ce cas, la réponse simple est de calculer la proportion de réponses identiques entre les deux juges. Par ailleurs, il est aussi possible d'utiliser la proportion d'accord, la proportion d'accords spécifiques, la proportion corrigée d'accord, etc. Lorsque le nombre de juges augmente, l'accord peut être défini de plusieurs façons. Par exemple, on peut considérer que plusieurs juges sont d'accord si toutes les réponses sont identiques ou si une majorité des réponses sont identiques. Bien que les divers coefficients aient été généralisés au cas de plusieurs juges, des formules permettant de déterminer un intervalle de confiance sont manquantes.

La seconde partie de la présentation de Vanbelle vise à illustrer les différences de couverture des intervalles de confiance des mesures d'accord en comparant la méthode Delta (δ) et la méthode établie par de Vet et al. (2017). Les propriétés statistiques des intervalles de confiance obtenus par ces différentes méthodes sont comparées sur des données simulées. La méthode δ est préférable, dans une vaste majorité des situations, à la méthode établie par de Vet et al. (2017). Un algorithme est proposé pour déterminer la taille d'échantillon lorsqu'aucun calcul analytique n'est possible. Ces méthodes sont disponibles dans une application Shiny.²

Vanbelle conclut sa présentation en soulignant l'importance de rapporter les différents indices d'accord accompagné d'un intervalle de confiance plutôt que d'utiliser une règle de classification servant de guide pour identifier un accord jugé acceptable, de calculer l'effectif d'échantillon nécessaire pour avoir des mesures fiables de l'accord interjuges (et intrajuges) et le besoin de rédiger des lignes directrices pour mieux aiguiller les chercheurs et chercheuses, mais également les praticiens et praticiennes qui doivent prendre des décisions éclairées sur le terrain.

Analyse factorielle par annulation du signal

La conférence d'André Achim (Université du Québec à Montréal) présente une approche novatrice, voire révolutionnaire, de l'analyse factorielle exploratoire (AFE). L'AFE requiert de prendre plusieurs décisions comme identifier le nombre de dimensions requises, décomposer la matrice de corrélation, traiter les variables orphelines et choisir le type et la méthode de rotation qui peuvent poser certains problèmes statistiques, comme les cas de Heywood et l'indétermination des facteurs doublets.

L'approche novatrice proposée par Achim, l'analyse factorielle par l'annulation du signal (SCFA), permet d'extraire les informations importantes sur la structure factorielle à découvrir, tout en réduisant le nombre de

décisions à prendre. SCFA postule uniquement que chaque facteur s'exprime dans au moins deux variables qui lui sont propres. L'idée sous-jacente est qu'un contraste sur une paire donnée de variables permet d'identifier un facteur commun. Pour chaque paire de variables, si une variable adéquatement pondérée permet d'annuler l'effet (le signal) d'une seconde variable, alors il y a potentiellement un facteur commun ; si l'annulation ne permet pas d'annuler le signal, alors a) la variable fait partie d'un autre facteur ou b) une des variables (ou les deux) est partagée sur plus d'un facteur.

L'annulation de signal est optimisée en minimisant la corrélation au carré maximale du contraste avec les variables restantes, ce qui procure un χ^2 avec $v - k$ degrés de liberté, où v est le nombre total de variables et k est le nombre de variables contrastées, et un poids optimisé w associé au contraste recherché. La première étape est de tester toutes les paires de variables afin d'obtenir un χ^2 et les pondérations w optimales pour chacune. Une analyse de groupement (*cluster analysis*) par dendrogramme est ensuite utilisée pour déterminer le nombre de facteurs à retenir à partir des χ^2 qui correspondent à des distances. De ce nombre, on soustrait les facteurs potentiels qui sont coplanaires avec deux autres facteurs; deux variables influencées proportionnellement par deux facteurs peuvent aussi annuler leur signal respectif et former un facteur potentiel à exclure. Une fois le nombre de facteurs potentiels identifié, SCFA teste l'annulation de signal sur des trios et ainsi de suite, pour identifier l'appartenance des variables restantes sur les facteurs. Enfin, les poids w permettent d'estimer la valeur de leurs saturations sur leur(s) facteur(s) respectifs ainsi que la matrice de corrélations interfacteurs.

Les avantages de cette approche sont multiples : aucune algèbre matricielle nécessaire, aucun problème de facteur doublet, aucun cas de Heywood, aucune indétermination de la rotation. Achim termine sa présentation en discutant de futurs développements de SCFA et en mentionnant que les versions Matlab et R sont disponibles sur Github.³

Processus d'élaboration et de validation d'un questionnaire francophone sur les compétences socioémotionnelles des enseignants

La conférence d'Ibtissem Ben Alaya (Université du Québec à Montréal) présente les différentes étapes nécessaires pour procéder au processus d'élaboration et validation d'un questionnaire. Guidé par le manque de disponibilité et de validité des questionnaires sur les compétences socioémotionnelles des enseignants et enseignantes, le

²Application Shiny: svanbelle.shinyapps.io/simpleagree/ par Sophie Vanbelle.

³Matlab : github.com/AndreAchim/SCFA (par André Achim) ; R github.com/quantmeth/SCFA (par P.-O. Caron).



processus suivi par Ben Alaya et ses collègues afin de procéder à la validation de leur questionnaire se décrit en sept principales étapes (Frenette et al., 2019). La première étape consiste à déterminer le concept à l'étude afin d'accumuler des preuves de validité de contenu pour définir les compétences socioémotionnelles des enseignants et choisir le cadre conceptuel approprié. Une revue de la littérature permet d'identifier les cinq domaines de compétences socioémotionnelles (Collaborative for Academic, Social, and Emotional Learning (CASEL), 2020) : la conscience de soi, l'autorégulation, la conscience sociale, les compétences relationnelles et la prise de décision responsable. Ces différents domaines peuvent se ventiler en deux volets : intrapersonnel (réfère à soi-même, ses propres pensées et émotions) et interpersonnel (implique l'autre). Pour la deuxième étape, l'équipe détermine le contexte de passation du questionnaire qui, dans le contexte pandémique, s'avère limité à des questionnaires administrés en ligne. Pour la troisième étape, l'équipe génère un bassin d'énoncés, une étape importante permettant d'appuyer la validité de contenu, en traduisant de l'anglais une liste d'énoncés issus de trois échelles identifiées dans la littérature scientifique, et en élimine les énoncés redondants. En quatrième étape, l'équipe détermine l'échelle de réponse en consultant des experts dans le domaine de l'éducation et des experts du milieu. La cinquième étape consiste à administrer le bassin d'énoncés final à un premier comité composé de professeurs, professeuses et professionnels, professionnelles dans le domaine de l'éducation afin de juger de la qualité des énoncés (degré de précision, pertinence, concordance avec le cadre conceptuel), de proposer des modifications et d'identifier les meilleures échelles de réponse. Suite aux modifications et retraits suggérés par ces experts et expertes, l'équipe a administré le questionnaire à un deuxième comité formé d'intervenants et d'intervenantes du milieu scolaire et des enseignants et enseignantes (de chaque niveau scolaire) et ce, en utilisant les mêmes critères que le premier groupe d'experts et d'expertes. Après avoir apporté ces dernières modifications et s'être assurée que leurs énoncés couvraient bien les cinq domaines définis plus tôt, comme sixième étape, l'équipe administre un prétest à des enseignants et enseignantes. La septième étape consiste à analyser les données recueillies lors du prétest et de la collecte finale. Les analyses effectuées sont la cohérence interne (mesurée par le α de Cronbach avec seuil mini-

mal de .70), les corrections items-total corrigées (avec un seuil minimal de .30), les corrélations interitems, l'analyse factorielle en axes principaux et l'analyse factorielle confirmatoire (seulement pour les données de la collecte finale). L'équipe a ainsi pu combler un vide dans la littérature scientifique francophone en élaborant un questionnaire visant à examiner les compétences socioémotionnelles des enseignants et enseignantes.

Contre le contrôle alpha : Les pratiques

En guise de continuité à des discussions ayant pris part lors de l'édition précédente de MQSH, Louis Laurencelle (Université du Québec à Trois-Rivières) discute du contrôle global α , soit le contrôle de l'erreur de Type I. D'entrée de jeu, Laurencelle mentionne que les chercheurs et chercheuses, qu'ils procèdent par enquêtes, par expérimentation en laboratoire ou en compilant des banques numériques, sont depuis quelques décennies, acculés à des méthodes de preuve statistique inconsistantes et très sévères. Les critères de correction du α , comme la correction de Bonferroni ou la différence significative honnête de Tukey, pilonnent

La 13e édition du colloque MQSH fut marquée par des discussions animées et permettant de porter une réflexion plus critique sur les enjeux actuels en méthodes quantitatives.

la puissance heuristique et statistique de leurs travaux de recherche. Sur le α global ou le contrôle global de α , c'est à travers les différentes pratiques mises de l'avant par leurs proposeurs que Laurencelle examine et argumente la nocivité d'une telle approche et de son absurdité.

Conclusion

La 13e édition du colloque MQSH fut marquée par des discussions animées et permettant de porter une réflexion plus critique sur les enjeux actuels en méthodes quantitatives. Les conférenciers ont pu bénéficier de l'expertise de chacun afin de suggérer des pistes d'exploration futures. La journée s'est close sur une discussion concernant l'avenir du MQSH, notamment en ce qui concerne la centralisation des informations du colloque. L'objectif proposé par Pier-Olivier Caron est d'offrir une stabilité au MQSH, par la création du site web où les informations des dernières éditions sont présentes.⁴ Un appel aux suggestions de thème a été énoncé pour la prochaine édition. Enfin, la 14e édition du MQSH sera organisée par Sébastien Béland et se tiendra à l'Université de Montréal.

Notes des auteurs

Olivier Vivier et Eva Lefebvre ont contribué également au présent compte rendu. Les auteur.e.s André Achim à

⁴Site web du colloque : mqsh.telug.ca.



Sophie Vanbelle sont en ordre alphabétique. Pier-Olivier Caron tient à remercier l'Université TÉLUQ pour le financement lié à l'évènement.

References

- Caron, P.-O., Bellemare-Lepage, A., Casgrain, J., Achim, A., Béland, S., Cousineau, D., Frenette, E., Lacourse, É., Laliberté, F., & Laurencelle, L. (2023). Compte rendu du colloque "méthodes quantitatives en sciences humaines (MQSH) 2023". *The Quantitative Methods for Psychology*, 19(4), 296–301. doi: [10.20982/tqmp.19.4.p296](https://doi.org/10.20982/tqmp.19.4.p296).
- Collaborative for Academic, Social, and Emotional Learning (CASEL). (2020). *CASEL'S SEL Framework: What are the core competence areas and where are they promoted?* Retrieved July 6, 2024, from <https://casel.org/wp-content/uploads/2019/12/CASEL-Competencies.pdf>
- de Vet, H. C., Dikmans, R. R., & Eekhout, I. (2017). Specific agreement on dichotomous outcomes can be calculated for more than two raters. *Journal of Clinical Epidemiology*, 88, 85–89. doi: [10.1016/j.jclinepi.2016.12.007](https://doi.org/10.1016/j.jclinepi.2016.12.007).
- Frenette, E., Fontaine, S., Hébert, M.-H., & Éthier, M. (2019). Étude sur la propension à tricher aux examens à l'université : Élaboration et processus de validation du questionnaire sur la tricherie aux examens à l'université (qteu). *Mesure et évaluation en éducation*, 42(2), 1–33. doi: [10.7202/1071514ar](https://doi.org/10.7202/1071514ar).
- Hernandez Engelhart, C., Gundro Brurberg, K., Aanstad, K. J., Pay, A. S. D., Kaasen, A., Blix, E., & Vanbelle, S. (2023). Reliability and agreement in intrapartum fetal heart rate monitoring interpretation: A systematic review. *Acta Obstetrica et Gynecologica Scandinavica*, 102(8), 970–985. doi: [10.1111/aogs.14591](https://doi.org/10.1111/aogs.14591).
- Holzinger, K. J. (1935). *Preliminary report on spearman-holzinger unitary trait study (no 5). introduction to bifactor theory: Solid and hollow staircase patterns for sets of data from mooseheart*. Department of Education.
- Spearman, C. (1927). *The abilities of man*. Macmillan.
- Vanbelle, S. (2024). Statistical inference for agreement between multiple raters on a binary scale. *British Journal of Mathematical and Statistical Psychology*, 77, 245–260. doi: [10.1111/bmsp.12333](https://doi.org/10.1111/bmsp.12333).

English abstract

On Friday, June 7, 2024, the 13th Annual meeting on Quantitative Methods in Human Sciences (MQSH) was held at TÉLUQ University (Montréal). Five speakers presented their work. Eric Frenette discussed the evolution of the bifactor model, from its proposal by Holzinger following Spearman's works in 1927 to its contemporary version, akin to Frankenstein's monster, having transformed its underlying assumptions and patched parameters to enhance its explanatory potential. Sophie Vanbelle, the invited speaker, presented the main statistical measures of inter-rater reliability and their interpretation to determine reliability and agreement for binary scales. She discussed the importance of reporting different agreement indices, calculating the necessary sample size to obtain reliable measures of inter-rater (and intra-rater) agreement, and the need for guidelines for researchers. André Achim proposed an innovative approach to exploratory factor analysis based not on matrix decomposition but on signal cancellation, which allows for extracting all relevant information, the number of factors, and the factor structure, while reducing the number of decisions. Ibtissem Ben Alaya presented the seven-step validation process of a questionnaire on teachers' socio-emotional skills, discussing the challenges of its implementation and the psychometric choices. Louis Laurencelle, in his presentation, counter-argued the various rationales and criteria for altering the classic significance alpha threshold, such as the Bonferroni or Tukey corrections, which undermine the statistical power of research endeavors in human sciences. He examined and argued the absurdity and harmfulness of such approaches. In conclusion, this 13th edition of the MQSH meeting lived up to its banner: competence, critique, and creativity.

Citation

Caron, P.-O., Vivier, O., Lefebvre, E., Achim, A., Ben Alaya, I., Frenette, E., Laurencelle, L., & Vanbelle, S. (2024). Compte rendu du colloque "Méthodes Quantitatives en Sciences Humaines (MQSH) 2024". *The Quantitative Methods for Psychology*, 20(2), 156–160. doi: [10.20982/tqmp.20.2.p156](https://doi.org/10.20982/tqmp.20.2.p156).

Copyright © 2024, Caron et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Received: 05/07/2024