
La recherche expérimentale sur l'efficacité du redoublement : entre éthique de l'expérimentation humaine et éthique méthodologique

Résumé

L'avènement de la méta-science a provoqué un séisme dans les disciplines qu'elle visait originellement à synthétiser. Ce que d'aucuns appellent désormais " la crise de la répliquabilité " (Ioannidis, 2005) a résonné dans de nombreux champs de recherche, dont celui des sciences de l'éducation (Plucker & Makel, 2021), en mettant notamment en évidence la faible qualité méthodologique de la recherche.

La présente communication se fonde sur deux méthodologies propres à la méta-science – la *Methodological Overview of Reviews* et la revue systématique de la littérature – afin de décrire l'évolution de la qualité méthodologique des études expérimentales ($k = 115$) et des méta-analyses ($k = 9$) menées quant aux effets du redoublement. L'analyse diachronique de la qualité de ces études permet, entre autres, de mettre en évidence une augmentation de la taille des effectifs concernés par les études. Si ces éléments paraissent scientifiquement réjouissants, cette évolution s'est produite en parallèle de concessions méthodologiques : le recours à des designs causaux alternatifs (*RDD*, *DiD*) dont la validité fait débat (Cook, 2002), le recours à des bases de données nationales longitudinales impliquant des comparaisons de groupes à effectifs inégaux ainsi que le recours à des stratégies de sélection *post-hoc* des élèves comparés. Cette augmentation de la taille des effectifs allant de pair avec une forme d'accommodation vis-à-vis de certaines concessions méthodologiques illustre la mise en concurrence actuelle de deux logiques distinctes de l'éthique de la recherche : l'éthique de l'expérimentation humaine, qui se fonde sur le principe du *Primum non nocere* (Salter, 1990) et l'éthique de la rigueur méthodologique, qui implique l'étude des effets d'une intervention sur des effectifs quantitativement suffisants pour atteindre des seuils minimaux de puissance statistique et de précision dans l'estimation des paramètres (Maxwell & Kelley, 2011).

A cet égard, les méta-analyses publiées jusqu'ici quant aux effets du redoublement indiquent que les effets attendus de cette pratique semblent, majoritairement, délétères, mais d'une ampleur assez faible (Allen et al., 2009 ; Goos et al., 2021). Cette apparente faiblesse de l'ampleur de l'effet attendue implique de recourir à des effectifs volumineux afin d'atteindre une puissance statistique satisfaisante (Barker Bausell & Li, 2002). D'ailleurs, notre revue systématique a montré que les études expérimentales menées quant aux effets du redoublement sont sous-puissantes. Du point de vue des politiques de recherche, la définition d'une position épistémique devient alors une véritable gageüre, puisque les deux logiques éthiques mènent à une impasse : doit-on continuer à massivement faire redoubler afin d'être en mesure d'étudier les effets de cette pratique au travers de recherches suffisamment puissantes ? Ou doit-on encourager la fin d'une pratique dont l'inefficacité semble démontrée depuis longtemps (Holmes, 1989 ; Jimerson, 2001), sans pour autant bénéficier d'une étude suffisamment puissante permettant de s'assurer des effets décrits par le passé ?

Ainsi, la présente communication propose une hypothèse explicative inédite de la crise de la " répliquabilité " en sciences de l'éducation à l'aune de ces contraintes paradoxales imposées aux

chercheurs par le biais des deux logiques éthiques concurrentes. En ce sens, cette communication entre dans l’Axe 3 du colloque (” Validation des résultats dans différents contextes ”) puisque ces contraintes éthiques semblent avoir un impact certain sur les méthodologies de recherche utilisées par les chercheurs et, partant, sur les résultats qui en découlent.

Si les concessions méthodologiques actuelles peuvent d’ailleurs sembler critiquables, la finalité du présent travail méta-scientifique est de définir les perspectives méthodologiques de la recherche de demain en se fondant sur ce que nous apprend, y compris par ses défauts, la recherche d’hier.

Bibliographie

Allen, C. S., Chen, Q., Wilson, V. L., & Hughes, J. N. (2009). Quality of research design moderates effects of grade retention on achievement: a meta-analytic, multilevel analysis. *Educational Evaluation and Policy Analysis, 31*, 480-499. <https://doi.org/10.3102/0162373709352239>

Barker Bausell, R., & Li, Y.-F. (2002). *Power analysis for experimental research : a practical guide for the biological, medical and social sciences*. Cambridge University Press

Cook, T. D. (2002). Randomized experiments in educational policy research: a critical examination of the reasons the educational evaluation community has offered for not doing them. *Educational Evaluation and Policy Analysis, 24*, 175-199. <https://doi.org/10.3102/01623737024003175>

Goos, M., Pipa, J., & Peixoto, F. (2021). Effectiveness of grade retention: a systematic review and meta-analysis. *Educational Research Review, 34* (100401), 1-14. <https://doi.org/10.1016/j.edurev.2021.100401>

Holmes, C. T. (1989). Grade level retention effects: a meta-analysis of research studies. In L. A. Shepard and M. L. Smith (eds.), *Flunking grades: research and policies on retention?* (pp. 16-33). The Falmer Press.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLOS Medicine 2*(8): e124. <https://doi.org/10.1371/journal.pmed.0020124>

Jimerson, S. R. (2001). Meta-analysis of grade retention research: implications for practice in the 21st century. *School Psychology Review, 30*, 420-437. Retrieved from: [http://web.a.ebscohost.com/ehost/detail, 734a-4c84-87ea-e42d8069bf5b%40sessionmgr4007&bdata=JnNpdGU9ZWwhvc3QtG12ZQ%3d%3d#AN=5508140&](http://web.a.ebscohost.com/ehost/detail?734a-4c84-87ea-e42d8069bf5b%40sessionmgr4007&bdata=JnNpdGU9ZWwhvc3QtG12ZQ%3d%3d#AN=5508140&)

Maxwell, S. E., & Kelley, K. (2011). Ethics and sample size planning. In A. T. Panter and S. K. Sterba (eds.), *Handbook of Ethics in Quantitative Methodology* (pp. 159-184). Routledge (Coll. Multivariate Applications Series).

Plucker, J. A., & Makel, M. C. (2021). Replication is important for educational psychology: Recent developments and key issues. *Educational Psychologist, 56*(2), 90-100. <https://doi.org/10.1080/00461520.2021.1895796>

Salter, D. C. (1990). Ethics of human testing. *International Journal of Cosmetic Science, 12*(4), 165-173. <https://doi.org/10.1111/j.1467-2494.1990.tb00532.x>

Mots-Clés: Méta, Science, Ethique de la recherche, Qualité de la recherche, Redoublement, Puissance statistique