

La politique fondée sur les preuves à l'épreuve du terrain

Le cas des méta-analyses sur le redoublement

Caroline De Pascale¹,
Hugues Draelants² & Jo B. Helgetun³

Introduction

De l'émergence d'une nouvelle méthode d'analyse à sa popularisation

Depuis sa labellisation il y a une cinquantaine d'années par Gene Glass en 1976, la méta-analyse, cette méthode qui consiste « à rassembler les résultats empiriques d'études singulières pour en faire une synthèse reproductible et quantifiée »⁴, n'a cessé de gagner en attractivité. Cette méthode a d'abord cru en popularité dans le champ scientifique, car elle a été présentée comme une solution à l'inflation de la production scientifique et au problème pratique posé par l'abondance d'informations dans la recherche⁵. Si c'est en médecine qu'elle a acquis son statut actuel de « norme d'excellence » (*gold standard*) en matière de recherche⁶, elle s'est répandue dans de nombreux domaines où le label fonctionne comme un « signe extérieur de scientificité »⁷. Une évolution qui a été soutenue par la

1. Groupe interdisciplinaire de Recherche sur la Socialisation, l'Éducation et la Formation (GIRSEF), UCLouvain, caroline.depascale@uclouvain.be.
2. Groupe interdisciplinaire de Recherche sur la Socialisation, l'Éducation et la Formation (GIRSEF), hugues.draelants@uclouvain.be.
3. Groupe interdisciplinaire de Recherche sur la Socialisation, l'Éducation et la Formation (GIRSEF), jo.helgetun@uclouvain.be.
4. Patrice Laroche, *La méta-analyse : Méthodes et applications en sciences sociales*, Louvain-la-Neuve, De Boeck, 2015, p.7.
5. Voir e.a. Gene V. Glass, « Primary, Secondary, and Meta-Analysis of Research », *Educational Researcher*, vol. 5, n°10, 1976, p. 3-8; MeowLan Evelyn Chan, Richard D. Arvey, « Meta-Analysis and the Development of Knowledge », *Perspectives on Psychological Science*, vol. 7, n°1, 2012, p. 79-92; Jessica Gurevitch, Julia Koricheva, Shinichi Nakagawa, Gavin Stewart, « Meta-analysis and the science of research synthesis », *Nature*, vol. 555, n°7695, 2018, p. 175-182.
6. Michael Borenstein, Larry V. Hedges, Julian P.T. Higgins, Hannah R. Rothstein, *Introduction to Meta-Analysis*. Chichester, John Wiley and Sons, 2009 ; Matthias Egger, George Davey Smith, « Meta-analysis: Potentials and promise », *BMJ*, n°315, 1997, p.1371.
7. Pierre Bourdieu, *Homo Academicus*, Paris, Minit, 1984.

propagation de l'usage politique du savoir scientifique⁸. En effet, dans le champ politique, le courant de l'*evidence-based policy* (ou « politique fondée sur les preuves ») perçoit la méta-analyse comme un outil essentiel pour intégrer et combiner de vastes quantités de recherches afin de parvenir à des conclusions claires et utilisables. La méta-analyse est dès lors valorisée pour sa capacité à fournir des preuves solides et fiables afin d'étayer les décisions politiques et les pratiques professionnelles sur ce qui est censé fonctionner le mieux⁹. En situation de déficit de légitimité, la méta-analyse devient donc un outil attractif pour renforcer la légitimité de l'action publique¹⁰.

Son succès s'est notamment étendu à l'éducation, domaine où les publications de méta-analyses progressent plus rapidement que toutes les autres publications. On comprend aussi que l'attrait du politique ne repose pas seulement sur la qualité scientifique de cette méthode, mais procède aussi d'une forme de dépolitisation et de naturalisation de l'éducation, tendances auxquelles contribuent également la promotion des « essais contrôlés randomisés » en économie de l'éducation¹¹ ou encore le recours croissant aux sciences cognitives¹². Or la prudence s'impose lorsque la méta-analyse est utilisée dans les sciences humaines et sociales, comme l'éducation, qui sont des sciences historiques. C'est-à-dire des sciences qui se caractérisent par l'absence de « répétition des phénomènes se déroulant à contexte constant ou indifférent » et qui s'appuient dès lors sur des preuves « toujours tributaires d'un contexte singulier d'observation, de mesure et d'argumentation »¹³.

L'objet de cet article est de produire une double critique de la méta-analyse. La première critique est interne au champ scientifique, car la réputation de scientificité et même de *nec plus ultra* de la preuve dont jouit de plus en plus la méta-analyse est imméritée en raison de spécificités épistémologiques. Notre seconde critique

8. Peter Weingart, « Scientific expertise and political accountability: paradoxes of science in politics », *Science and Public Policy*, vol. 26, n°3, 1999, p. 151-161.
9. Ben Goldacre, *Building Evidence into Education*, Technical Report, 2013, p. 1-19 ; Ian Menter, « From interesting times to critical times? Teacher education and educational research in England », *Research in Teacher Education*, vol. 3, n°1, 2013, p. 38-40 ; Romuald Normand, *The Changing Epistemic Governance of European Education: The Fabrication of the Homo Academicus Europeanus?*, Berlin, Springer, 2017.
10. Hugues Draelants, Sonia Revaz, *L'évidence des faits. La politique des preuves en éducation*, Paris, Presses universitaires de France, 2022.
11. Arthur Jatteau, « Les essais contrôlés randomisés. Une comparaison entre la médecine et l'économie », *Philosophia Scientiæ. Travaux d'histoire et de philosophie des sciences*, vol. 23, n°2, 2019, p. 85-110.
12. Bernard Lahire, Claude Rosental, *La cognition au prisme des sciences sociales*, Paris, Archives contemporaines, 2008.
13. Jean-Claude Passeron, « La forme des preuves dans les sciences historiques », *Revue européenne des sciences sociales*, vol. XXXIX, n°1, 2001, p. 42.

est sociologique, ou sociopolitique. Elle soutient que le politique fait un usage excessif de cette forme de preuve qui ne tient pas ses promesses et suscite un faux sentiment de certitude : ce qui est censé être efficace (la rhétorique du « *what works* ») ne l'est guère, en fin de compte, pour fonder la légitimité de l'action publique. Pour soutenir ce point, nous allons nous appuyer sur une étude de cas consacrée à la question du redoublement. Les recherches sur le redoublement ont en effet été fortement marquées par plusieurs méta-analyses, l'étude la plus citée sur le redoublement étant d'ailleurs l'une d'entre elles¹⁴. En outre, les politiques qui visent à réguler le redoublement auxquelles nous nous sommes intéressés sont sous-tendues par une expertise fortement adossée à ces méta-analyses.

Avant d'en venir au cœur de notre propos, nous commencerons par expliquer ce qu'est une méta-analyse et par préciser la place que cette méthode occupe dans le champ de la recherche en éducation, plus particulièrement, dans le domaine des recherches sur le redoublement.

Éléments de contexte

Qu'est-ce qu'une méta-analyse ?

La méthode de synthétisation d'études quantitatives portant sur un objet de recherche similaire est bien antérieure à la labellisation proposée par Glass en 1976¹⁵. Il y a cent cinquante ans, les astronomes s'interrogeaient déjà sur la manière de sélectionner puis de combiner des observations de mesures physiques de faible qualité¹⁶. Ceci a abouti à l'élaboration d'une théorie générale à ce sujet¹⁷. Deux décennies plus tard, en épidémiologie, Pearson combinait différentes séries de données sur la fièvre entérique, dont il estimait la taille trop petite pour élaborer des conclusions si elles étaient prises individuellement¹⁸. Dans les années 1930, c'est au travers de questions d'optimisation agricoles¹⁹ qu'une série

14. Shane Jimerson, « Meta-analysis of Grade Retention Research: Implications for Practice in the 21st Century », *School Psychology Review*, vol. 30, 2001, p. 420-437

15. Gene V. Glass, « Primary, Secondary, and Meta-Analysis of Research », art. cit.

16. Benjamin Peirce, « Criterion for the rejection of doubtful observations », *Astronomical Journal*, vol. 2, n°4, 1852, p. 161-163 ; George Biddell Airy, *On the Algebraical and Numerical Theory of Errors of Observations and the Combination of Observations*, Londres, Macmillan and Co., 1875.

17. Simon Newcomb, « A Generalized Theory of the Combination of Observations so as to Obtain the Best Result », *American Journal of Mathematics*, vol. 8, n°4, 1886, p. 343-366.

18. Karl Pearson, « Report on Certain Enteric Fever Inoculation Statistics », *British Medical Journal*, vol. 2, n°2288, 1904, p. 1243-1246.

19. George Dunlop, « Methods of Experimentation in Animal Nutrition », *The Journal of Agricultural Science*, vol. 23, n°4, 1933, p. 580-614 ; Jay L. Lush, « Predicting gains in feeder cattle and pigs », *Journal of Agricultural Research*, vol. 42, 1931, p. 853-881.

d'ouvrages voient le jour pour synthétiser les résultats de divers travaux²⁰, en même temps que les statisticiens poursuivent leurs recherches sur ce thème²¹. Au milieu des années 1960, l'enseignement des techniques de synthèse sous-tendant la méta-analyse a déjà lieu dans les universités²². En conséquence de quoi, le point de bascule de l'article de Glass, n'est donc pas tant de développer une méthode d'analyse inédite, que de la nommer et d'en dessiner les contours. Ces contours quels sont-ils ? Glass apporte l'éclairage suivant :

« Cinq cents études sur la taille des classes ou le regroupement des aptitudes peuvent s'accumuler ; elles défieront tout résumé. [...] Le fait d'intégrer les résultats de la recherche en éliminant les études "mal faites" revient à écarter une grande quantité de données importantes. Leur signification ne peut pas plus être saisie [...] sans l'aide de techniques d'organisation, de représentation et d'interrelation des données. »²³

En somme, la méta-analyse se fonde sur deux éléments principaux : elle peut synthétiser les résultats d'études sur un sujet, nonobstant leur qualité disparate, et elle apporte un surcroît de savoir au regard des revues narratives de littérature, sous la forme d'un effet global quantifié. Les premiers travaux des méta-analyses dans le domaine de la psychologie ont alors visé à prouver que la psychothérapie fonctionnait²⁴.

C'est d'ailleurs précisément en psychologie et en médecine que la méta-analyse a pris une ampleur considérable, notamment du fait de « *la dimension fondamentalement expérimentale de ces études, d'une part, et [de] l'enjeu que représentent leurs conclusions pour les praticiens de la santé* »²⁵. Dans un contexte d'inflation de la production scientifique, et compte tenu de cet enjeu, il est nécessaire de pouvoir synthétiser ces études. D'autres auteurs pointent

20. William G. Cochran, « Problems Arising in the Analysis of a Series of Similar Experiments », *Supplement to the Journal of the Royal Statistical Society*, vol. 4, N°1, 1937, p.102-118 ; George W. Snedecor, *Statistical Methods applied to experiments in Agriculture and Biology*, Ames, Collegiate Press, 1937.
21. Ronald A. Fisher, *Statistical methods for research workers*, 5th Edition, Londres, Oliver & Boyd, 1934 ; Karl Pearson, « On a Method of Determining Whether a Sample of Size N Supposed to Have Been Drawn from a Parent Population Having a Known Probability Integral has Probably Been Drawn at Random », *Biometrika*, vol. 25, N°3-4, 1933, p. 379-410.
22. Robert Rosenthal, *Meta-Analytic Procedures for Social Science Research*, Beverly Hills, Sage Publications, 1984.
23. Gene V. Glass, « Primary, Secondary, and Meta-Analysis of Research », art. cit., p. 4.
24. Robert Rosenthal, *Meta-Analytic Procedures for Social Science Research*, op. cit. ; Mary Lee Smith, Gene V. Glass, « Meta-analysis of psychotherapy outcome studies », *The American psychologist*, vol. 32, N°9, 1977, p. 752-760.
25. Patrice Laroche, *La méta-analyse*, op. cit., p. 9.

la force de la méta-analyse pour fabriquer du consensus scientifique²⁶. Cette popularisation dans ces domaines et dans la vague de l'*Evidence Based Medicine* (EBM), qui fait de la preuve statistique et des synthèses d'essais contrôlés randomisés (ECR) le sommet de la hiérarchie de la preuve, a permis à la méta-analyse d'acquiescer ses lettres de noblesse et son statut privilégié²⁷. Or, dès le début, l'usage de la méta-analyse dans les sciences de l'éducation ne s'est pas focalisé sur les seuls ECR. Ainsi, l'une des premières méta-analyses visant à attester de l'effet de la taille des classes sur les attitudes et les apprentissages englobe diverses techniques quantitatives²⁸.

En conséquence, la labellisation proposée par Glass recouvre une diversité de sens à cause de ses utilisations variables selon les contextes²⁹, sans que l'appréciation globale portée à la méthode ne varie selon ces différents usages. L'une des raisons évoquées est que malgré les difficultés d'agrégation des études de différentes méthodes, la synthèse opérée par la méta-analyse reste meilleure que d'autres formes de synthèse pour assurer l'impartialité et la reproductibilité des résultats³⁰. Cependant, il n'est pas évident que les qualités attribuées à la méta-analyse en médecine soient transposables dans les sciences de l'éducation et de la formation.

Place des méta-analyses dans les sciences de l'éducation et dans la recherche sur le redoublement

En termes absolus, les méta-analyses ne représentent qu'une part marginale des publications en éducation, y compris dans les articles quantitatifs. Les méthodes de recherche les plus courantes qui y sont identifiées étant les méthodes causales-comparatives et corrélationnelles³¹. Les méta-analyses n'occupent donc

26. MeowLan Evelyn Chan, Richard D. Arvey, « Meta-Analysis and the Development of Knowledge », art. cit.
27. Hugues Draelants, Sonia Revaz, *L'évidence des faits*, op. cit.
28. Mary Lee Smith, Gene V. Glass, « Meta-analysis of Research on Class Size and Its Relationship to Attitudes and Instruction », *American Educational Research Journal*, vol. 17, n°4, 1980, p. 419-433.
29. Jean-Claude Barbier, « La précarité, une catégorie française à l'épreuve de la comparaison internationale », *Revue française de sociologie*, vol. 46, n°2, 2005, p. 351-371.
30. MeowLan Evelyn Chan, Richard D. Arvey, « Meta-Analysis and the Development of Knowledge », art. cit.
31. Sur les 100 publications les plus citées en psychologie de l'éducation, 70 avaient un design quantitatif, dont 35 utilisaient une méthode causale-comparative, 24 une méthode corrélationnelle, il y avait 6 méta-analyses. Patrick J. McNicholas, Randy G. Floyd, Lauren E. Fennimore, Sequoya A. Fitzpatrick, « Determining journal article citation classics in school psychology: An updated bibliometric analysis using Google Scholar, Scopus, and Web of Science », *Journal of School Psychology*, vol. 90, 2022, p. 94-113.

pas une position hégémonique dans les sciences de l'éducation telles qu'elles se pratiquent, néanmoins elles dominent symboliquement puisqu'elles sont considérées comme offrant le meilleur niveau de preuve dans le paradigme de l'*evidence-based*. Par ailleurs, si l'on se fonde sur les taux de progression des méta-analyses (voir tableaux ci-dessous), leur montée en puissance est indéniable³², y compris en éducation où le recours à la méta-analyse augmente plus vite que le reste des publications (Figures 1 et 2). Aujourd'hui on fait même des méta-analyses de méta-analyses³³.

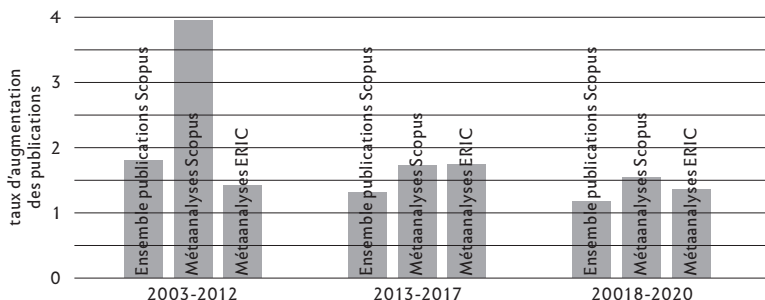


Figure 1. Comparaison des taux d'augmentation des publications et des méta-analyses dans le domaine de l'éducation (base de données Scopus, ERIC)

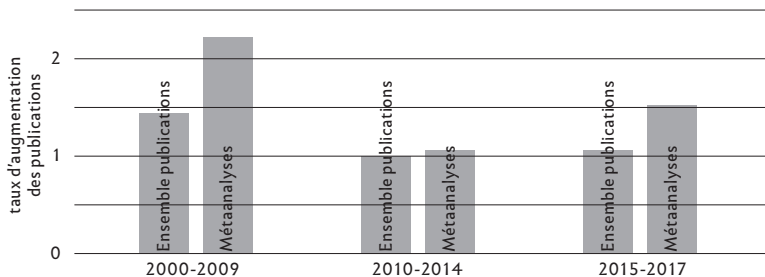


Figure 2. Comparaison des taux d'augmentation des publications et des méta-analyses dans le domaine de l'éducation (base de données ERIC Proquest)

32. Jessica Gurevitch, Julia Koricheva, Shinichi Nakagawa et Gavin Stewart, « Meta-analysis and the science of research synthesis », art. cit.
33. Cet élan vers la massification n'a pas été sans conséquence, en témoigne la rapidité avec laquelle certaines méga-analyses ont été critiquées voire qualifiées de pseudoscience. Voir à ce sujet Pierre-Jérôme Bergeron, « Comment faire de la pseudoscience avec des données réelles : une critique des arguments statistiques de John Hattie dans *Visible Learning* par un statisticien », *The MJE Forum*, vol. 51, n°2, 2016, p. 935-946.

Cette augmentation de l'emploi des méta-analyses est à replacer dans la perspective de la médicalisation de l'éducation, que Tröhler décrit comme une forme avancée de technocratisation³⁴. Depuis le dix-neuvième siècle, les problèmes sociaux sont, avance-t-il, progressivement devenus des défis éducatifs. Dans la foulée, la culture technocratique a placé l'expert, plus que le professionnel, au centre de l'élaboration de solutions aux défis nouvellement placés devant l'institution scolaire. Peu à peu, à partir des années 1970, l'éducation est ensuite entrée dans un paradigme médical fondé sur un positivisme social qui ne vise pas à comprendre un système mais évalue statistiquement ce qu'il produit. Le programme PISA en est un bon exemple ; il n'offre pas une meilleure compréhension du système scolaire, mais des données sur son état. Le modèle est celui des essais cliniques. Ces derniers ne sont pas conçus pour découvrir les raisons de l'efficacité d'un médicament, mais pour prouver qu'il fonctionne. La quête des preuves statistiques remplace alors celle de la compréhension. Notons que ce mouvement n'est pas porté par les médecins ou les enseignants, puisque ce paradigme tend à diminuer leur expertise professionnelle³⁵.

L'évolution de la recherche sur le redoublement est un exemple très parlant de médicalisation et c'est pourquoi nous nous fondons dessus pour appuyer nos critiques et notre discussion sur l'usage des méta-analyses dans les sciences de l'éducation et de la formation. Celles-ci sont au nombre de cinq³⁶, mais leur *design* hérite significativement des recommandations méthodologiques d'une synthèse narrative sur le sujet de la mesure des effets du redoublement réalisée en 1975 par Gregg B. Jackson³⁷. En particulier car celui-ci choisit d'intégrer des études de qualité

34. Daniel Tröhler, « The medicalization of current educational research and its effects on education policy and school reforms », *Discourse Studies in the Cultural Politics of Education*, vol. 36, n°5, 2015, p. 749-764.

35. *Ibid.*

36. Dans l'ordre de parution : C. Thomas Holmes et Kenneth M. Matthews, « The Effects of Nonpromotion on Elementary and Junior High School Pupils: A Meta-Analysis », *Review of Educational Research*, vol. 54, n°2, 1984, p. 225-236 ; C. Thomas Holmes, « Grade-level retention effects: A meta-analysis of research studies », in Lorrie A. Sheppard, Mary Lee Smith (eds.), *Flunking grades: Research and policies on retention*, Philadelphie, The Falmer Press, 1989, p. 16-33 ; Shane Jimerson, « Meta-analysis of Grade Retention Research: Implications for Practice in the 21st Century », *School Psychology Review*, vol. 30, n°3, 2001, p. 420-437 ; Chiharu S. Allen, Qi Chen, Victor L. Willson et Jan N. Hughes, « Quality of Research Design Moderates Effects of Grade Retention on Achievement: A Meta-Analytic, Multilevel Analysis », *Educational Evaluation and Policy Analysis*, vol. 31, n°4, 2009, p. 480-499 ; Mieke Goos, Joana Pipa et Francisco Peixoto, « Effectiveness of grade retention: A systematic review and meta-analysis », *Educational Research Review*, vol. 34, 2021, p. 1-14.

37. Gregg B. Jackson, « The Research Evidence on the Effects of Grade Retention », *Review of Educational Research*, vol. 45, n°4, 1975, p. 613-635.

médiocre dans sa synthèse, ce que les trois premières méta-analyses³⁸ reproduiront, avant que deux travaux soulignent que ce choix renforce les résultats négatifs du redoublement³⁹. Jackson choisit également d'intégrer des travaux portant sur une période temporelle très longue, car remontant jusqu'aux années 1920. Et les méta-analyses continuent d'intégrer des travaux portant sur des contextes temporels différents. Enfin, Jackson construit un modèle d'évaluation du redoublement dans lequel la méthode quasi-expérimentale est jugée la plus adéquate : les élèves redoublants sont appariés à des élèves promus. La promotion est considérée comme un proxy neutre dans les analyses statistiques⁴⁰. Ces trois éléments sont essentiels aux critiques sociologiques que nous allons porter par la suite.

Quelle est la place des méta-analyses dans l'étude du redoublement ? Il faut noter qu'alors qu'il n'y a que 5 méta-analyses sur 278 publications portant sur le redoublement (soit 1,79%) des publications, 3 méta-analyses font partie des 10 publications les plus citées sur la question (voir Tableau 1).

Rang	Articles	Citations
1	Jimerson, S.R. (2001)	283
3	Holmes, C.T., Matthews, K.M. (1984)	186
6	Jackson, G.B. (1975) – Synthèse narrative	126
10	Allen, C.S., Chen, Q., Willson, V.L., Hughes, J.N. (2009)	116
61	Goos, M., Pipa, J., Peixoto, F. (2021)	27

Tableau 1 : Classement des méta-analyses selon leur nombre de citations⁴¹
 Source : Scopus, résultats obtenus au moyen de la recherche suivante :
 TITLE (« grade retention ») OR TITLE (« grade repetition »)
 OR TITLE (« Grade level retention ») OR TITLE (« nonpromotion »)
 [Recherche réalisée le 19 mars 2024]

38. C. Thomas Holmes et Kenneth M. Matthews, « The Effects of Nonpromotion on Elementary and Junior High School Pupils: A Meta-Analysis », art. cit. ; C. Thomas Holmes, « Grade-level retention effects: A meta-analysis of research studies », art. cit. ; Shane Jimerson, « Meta-analysis of Grade Retention Research: Implications for Practice in the 21st Century », art. cit.
39. Jon Lorence et Anthony Gary Dworkin, « Elementary grade retention in Texas and reading achievement among racial groups: 1994-2002 », *Review of Policy Research*, vol. 23, N°5, 2006, p. 999-1033 ; Chiharu S. Allen, Qi Chen, Victor L. Willson et Jan N. Hughes, « Quality of Research Design Moderates Effects of Grade Retention on Achievement: A Meta-Analytic, Multilevel Analysis », art. cit.
40. Pour une discussion plus approfondie des choix opérés par Jackson, voir Hugues Draelants, « Le redoublement est-il vraiment moins efficace que la promotion automatique ? », *Les Cahiers de recherche du GIRSEF*, N°113, 2018, ojs.uclouvain.be/index.php/cahiersgirsef/article/view/54123
41. La méta-analyse de Holmes, parue en 1989 dans un livre et non répertoriée sur Scopus, n'est pas incluse.

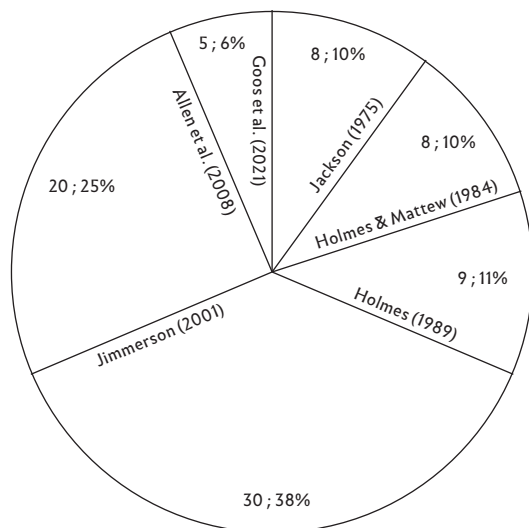


Figure 3 : Citations obtenues par les diverses méta-analyses sur le redoublement en 2021-2022

Des publications plus récentes ont pourtant discuté des lacunes techniques des travaux intégrés dans ces méta-analyses et des limites de leurs conclusions quant aux effets du redoublement⁴². Cependant, ces travaux plus récents n'ont pas une visibilité aussi importante que les méta-analyses plus anciennes dans le champ de la recherche. Sur les années 2021-2022, les quatre premières méta-analyses, pourtant critiquées, obtiennent encore 69 % des citations (figure 3). Ce poids est interpellant à plus d'un titre, notamment parce que 49 % des 41 publications publiées entre 2021 et 2022 qui se fondent sur ces méta-analyses ne citent aucune des méta-analyses plus récentes sur le redoublement. De plus, après lecture de ces travaux qui ne citent que les méta-analyses les plus anciennes, les trois quarts d'entre eux traitent les effets du redoublement comme si les résultats de la recherche étaient univoques. Au contraire, les publications qui se fondent sur les travaux plus

42. Karl L. Alexander, Doris R. Entwisle et Susan L. Dauber, *On the Success of Failure: A Reassessment of the Effects of Retention in the Primary School Grades*, Cambridge, Cambridge University Press, 2003 (2e éd.) ; Chiharu S. Allen, Qi Chen, Victor L. Willson et Jan N. Hughes, « Quality of Research Design Moderates Effects of Grade Retention on Achievement: A Meta-Analytic, Multilevel Analysis », art. cit. ; Robert Gary-Bobo, Jean-Marc Robin, « La question des redoublements. Analyse économique et problèmes statistiques », art. cit. ; Robert Gary-Bobo, Marion Goussé et Jean-Marc Robin, « Grade retention and unobserved heterogeneity », *Quantitative Economics*, vol. 7, n°3, 2016, p. 781-820 ; Jon Lorence, « Retention and Academic Achievement Research Revisited from a United States Perspective », *International education journal*, vol. 7, n°5, 2006, p. 731-777.

récents ont tendance ($r = 0.418^{**}$) à rapporter que les résultats en la matière sont équivoques et que les effets du redoublement varient selon les variables sélectionnées⁴³. Il est donc questionnant que ces travaux anciens continuent à avoir autant d'importance dans la recherche actuelle, alors qu'ils propagent des idées contestées sur la question.

Cette prévalence de travaux anciens et contestés qui dépeignent une image univoque et négative des effets du redoublement a pour conséquence de créer la (fausse) impression que la recherche sur le redoublement est unanime ou, pire, qu'il n'est plus nécessaire d'entreprendre de nouvelles études. Cependant, d'autres difficultés de nature épistémologique émanent également des méta-analyses plus récentes et à la méthodologie plus aboutie. Au vu de la place importante des méta-analyses dans la littérature actuelle sur le redoublement, il est nécessaire de s'interroger en détail sur leurs limites.

Les méta-analyses proposent-elles une bonne «abstraction du réel»?

Des difficultés épistémologiques liées à l'usage de la méta-analyse en général et pour l'étude du redoublement en particulier

La position de la méta-analyse au sommet de la pyramide des preuves ne l'a pas préservée des critiques épistémologiques. Très tôt, des chercheurs ont souligné les limites de l'emploi des méta-analyses en éducation⁴⁴. Souvent formulées par des usagers de la méthodes, ces discussions témoignent de leur réflexivité et de leur capacité à développer une critique endogène. Dès 1984, Slavin souligne que cette méthode a pris «un mauvais départ» dans les sciences de l'éducation⁴⁵. Il considère que la qualité des études qui composent les méta-analyses affecte la qualité de ces dernières.

43. Les 41 articles publiés sur le redoublement en 2021 et 2022 ont été codés en fonction de leur citation des méta-analyses selon qu'elles étaient récentes (à partir de 2009) ou anciennes. Nous avons aussi codé leur perspective sur les résultats portant sur le redoublement (univoque ou équivoque). Par ailleurs, il y a toujours 20 travaux (donc la moitié) qui ne citent que les MA plus anciennes. La corrélation a été calculée entre la date d'usage des méta-analyses et le fait de rapporter des résultats univoques ou équivoques.

44. Robert E. Slavin, «Meta-Analysis in Education: How Has It Been Used?», *Educational Researcher*, vol. 13, n°8, 1984, p. 6-16; James A. Kulik, Chen-Lin C. Kulik, «Meta-Analysis in Education», *International Journal of Educational Research*, vol. 13, n°8, 1989, p. 221-340.

45. Robert E. Slavin, *ibid.*, p. 13.

Plus précisément, cet auteur montre que des chercheurs noient des études de bonne qualité à taille d'effet moindre dans des travaux de qualité médiocre pourtant sur de grandes tailles d'effet. Le fait que des données recueillies par des procédés de piètre qualité aboutissent à des méta-analyses de qualité tout aussi discutable est un problème désormais bien identifié et désigné sous l'appellation du « *garbage in, garbage out* » ou du « *junk in, junk out* »⁴⁶. Pour garantir des résultats fiables, il convient donc d'évaluer la qualité méthodologique des études incluses. L'hétérogénéité entre les variables et les études incluses doit également être prise en compte sous peine de fausser les résultats⁴⁷. Pour donner un exemple simple, si on synthétise des études qui évaluent l'effet de l'alimentation sur la prise de poids, mais qu'une première étude inclut la présence d'écran lors des repas, tandis qu'une deuxième étude mesure par ailleurs l'activité physique, leur association dans la même méta-analyse produira des résultats incohérents, biaisés ou inexploitable.

Si la pertinence de recourir à la méta-analyse a fait l'objet de critiques en médecine, son usage *a fortiori* en éducation pose encore plus questions, et ce pour plusieurs raisons.

La place privilégiée des méta-analyses doit d'abord être interrogée dans les sciences de l'éducation et de la formation, car ces sciences ont la particularité d'être un ensemble composite de disciplines, relevant d'épistémologies différentes, ce dont cette méthode ne tient pas compte. Les sciences de l'éducation et de la formation sont marquées par deux tendances épistémologiques : d'une part une production de connaissance à orientation interdisciplinaire, et d'autre part l'établissement de savoirs scientifiques à partir de l'ancrage épistémologique d'une discipline établie⁴⁸. Comme nous l'avons mentionné, la méta-analyse est conçue pour la synthèse de travaux expérimentaux, qui sont particulièrement publiés en psychologie de l'éducation depuis les années 1980⁴⁹. Les méta-analyses en sciences de l'éducation donnent une place particulière à un savoir disciplinaire, ainsi qu'aux approches essentialisantes du réel (donc d'héritage positiviste) tout en ayant trait

46. John P.A. Ioannidis, « The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses ». *The Milbank quarterly*, vol. 94, N°3, 2016, p. 485-514.

47. Jessica Gurevitch, Julia Koricheva, Shinichi Nakagawa et Gavin Stewart, « Meta-analysis and the science of research synthesis », art. cit.

48. Brigitte Albero, *Manuel de sciences de l'éducation et de la formation*, op. cit.

49. Jean-Marie De Ketele, « La publication scientifique en sciences de l'éducation et de la formation : état des lieux, dangers et perspectives », *Les dossiers des sciences de l'éducation*, N°41, 2019, p. 43-59.

à des thématiques qui sont pourtant abordées simultanément par plusieurs disciplines. En outre, les productions scientifiques en éducation ont une visée pragmatique « *visant à documenter plus directement l'action (politique, ingénierie, didactique, pédagogique, technique, etc.)* »⁵⁰. Cette dimension praxéologique est aussi axiologique puisque la documentation des actions humaines en éducation est couplée à un discours souvent critique sur l'état des systèmes éducatifs. Cette double dimension se traduit par des recommandations pour l'action, notamment dans le courant de l'*evidence-based* : le savoir scientifique a pour vocation d'être transformé en action pratique et en action publique.

Les complexités épistémiques mentionnées ci-dessus sont donc centrales dans la recherche en éducation, en raison du statut hybride de l'éducation qui a ceci de particulier d'être à la fois un domaine de recherche et un champ de pratique. Ce qui confère aussi à l'éducation sa nature complexe c'est que de nombreuses variables dépendantes du contexte sont à l'œuvre et interagissent simultanément, quel que soit le thème étudié : de la politique de l'éducation⁵¹ aux théories de l'apprentissage/pédagogie⁵² ainsi qu'aux formes institutionnalisées de la connaissance⁵³. Concrètement, si l'on ne parvient pas à répliquer des études en laboratoire contrôlé, comme le montre la crise de réplication en psychologie⁵⁴, comment imaginer qu'elles fonctionneront dans des environnements non contrôlés tels qu'une salle de classe ?

Surtout, la question de ce qui marche le mieux n'est pas neutre (l'est-elle jamais ?), car cette question axiologique ne peut se départir des finalités des systèmes éducatifs⁵⁵. En éducation, retenir ce qui marche dépend des missions assignées à l'école et est donc une question politique. C'est pourquoi le fait que de très

50. Brigitte Albero, *Manuel de sciences de l'éducation et de la formation*, Louvain-la-Neuve, De Boeck, p. 68.
51. Stephen J. Ball, « What Is Policy? Texts, Trajectories and Toolboxes », *Discourse: Studies in the Cultural Politics of Education*, vol. 13, N°2, 1993, p. 10-17.
52. Karl L. Alexander, Doris R. Entwisle et Linda S. Olson, « Schools, Achievement, and Inequality: A Seasonal Perspective », *Educational Evaluation and Policy Analysis*, vol. 23, N°2, 2001, p. 171-191.
53. W. Richard Scott, *Institutions and organizations: ideas, interests and identities*, Londres, Sage, 2014 ; Jo Helgetun, Ian Menter, « From an age of measurement to an evidence era? Policy-making in teacher education in England », *Journal of Education Policy*, 2020, vol. 37, N°1, p. 1-18 ; Jo Helgetun, « The importance of context: Teacher education policy in England and France compared », in Ian Menter (ed.), *The Palgrave Handbook of Teacher Education Research*, Cham, Palgrave, 2022.
54. Stuart J. Ritchie, *Science Fictions: Exposing Fraud, Bias, Negligence and Hype in Science*, Londres, The Bodley Head, 2020.
55. Gert Biesta, « Why 'What Works' won't work: Evidence-based practice and the democratic deficit in educational research », *Educational Theory*, vol. 57, N°1, 2007, p. 1-22.

nombreuses méta-analyses se fondent sur des essais contrôlés randomisés devrait amener à prendre avec encore plus de précautions les résultats de ces études, au risque d'aggraver le problème de l'interventionnisme en éducation⁵⁶. Ce dernier, qui valorise le mesurable et devient par conséquent sujet à un argumentaire démonstratif du type « la science prouve », occulte la réflexion sur les finalités de l'éducation⁵⁷.

Une critique sociologique des études quantitatives sur le redoublement, synthétisées par les méta-analyses

Aux critiques épistémologiques évoquées plus haut, nous voudrions ajouter une série de critiques proprement sociologiques. La thématique du redoublement ne relève en effet pas seulement de l'efficacité des apprentissages (abordé par les disciplines de psychologie des apprentissages, de didactique et de sciences cognitives), mais aussi des inégalités scolaires et de l'éducation conçue comme un investissement et un capital, des questions abordées par les sciences sociales et la sociologie en particulier. Surtout, sociologiquement, appréhender le redoublement requiert une connaissance fine des processus et mécanismes sociaux qui façonnent le phénomène et conditionnent l'expérience qu'en ont les élèves et les enseignants, et ce qu'ils en attendent. L'analyse sociologique du redoublement invite aussi à resituer ce dispositif dans le fonctionnement d'ensemble du système éducatif, tout en tenant compte de la diversité des contextes scolaires. Ce faisant, elle donne à voir ce qui n'est pas nécessairement visible dans les approches positivistes et désincarnées du phénomène, qui s'avèrent en définitive assez préjudiciables lorsqu'elles deviennent la seule boussole de l'action publique, nous y reviendrons.

Notre première critique sociologique tient au critère de comparaison des élèves. Une partie significative des études quantitatives synthétisées par les méta-analyses compare les redoublants à des élèves de même âge⁵⁸. L'idée est ici qu'il existe une variation au

56. Cathy Burnett, Mike Coldwell, « Randomised controlled trials and the interventionisation of education », *Oxford Review of Education*, vol. 47, N°4, 2021, p. 423-438.
57. Juliane Lynch, Julie Rowlands, Trevor Gale et Stephen Parker, « An outline of a theory of practice methodologies: Education research as an expansive activist endeavor », in J. Lynch, J. Rowlands, T. Gale, et S. Parker (eds.), *Practice Methodologies in Education Research*, Londres, Routledge, 2020 ; Jo Helgetun, Ian Menter, « From an age of measurement to an evidence era? », art. cit.
58. Marcel Crahay, « Peut-on conclure à propos des effets du redoublement? », *Revue Française de Pédagogie*, 2004, N°148, p. 11-23.

niveau de la note scolaire et que certains élèves auraient pu obtenir des résultats suffisants pour passer dans une autre classe et une autre école. Ceci les rendrait comparables à des élèves ayant réussi avec de faibles résultats. Sur cette base, « *Les sujets d'un des groupes, que l'on qualifiera d'expérimental, bénéficient d'un traitement : le redoublement* »⁵⁹. Une autre partie des études compare les performances des élèves redoublants aux élèves de même niveau scolaire. Ces deux comparaisons sont majoritaires aujourd'hui, au point d'être les seules retenues dans la dernière méta-analyse⁶⁰.

Cependant ces comparaisons sont problématiques d'un point de vue sociologique. Il s'agit d'un simulacre d'expérimentation (on parle de modèle quasi-expérimental) qui, se faisant, désincarne l'élève de son parcours scolaire, de la signification de son redoublement, et le compare avec un autre élève qui n'est pas dans la même situation ni dans le même contexte scolaire. D'ailleurs, les enseignants privilégient la comparaison pré/post redoublement, choix qui est tout à fait praticable dans les études quantitatives⁶¹ et sociologiquement plus pertinent. La méthode d'appariement est en fait un choix épistémologique délibéré, basé sur l'idée que l'expérimentation est nécessaire. Cependant, cela n'évalue pas la progression de l'élève par rapport à ses performances ou son comportement l'année précédente. Dans un langage sociologique, il s'agit ici de comparer conjointement des trajectoires et des bifurcations⁶², dont les caractéristiques et la signification pour les élèves et leurs familles, ainsi que pour leurs professeurs n'est nullement similaire. Réaliser des comparaisons ayant peu de sens d'un point de vue sociologique risque de creuser l'écart entre l'évaluation des acteurs de terrain et celle des chercheurs. Une réforme politique fondée sur des études de ce type court alors le risque de se heurter à la réalité vécue par les praticiens, et de perdre en légitimité⁶³. Nous y reviendrons.

Deuxièmement, l'évaluation de l'efficacité du redoublement repose sur un critère de temporalité également questionnable d'un point de vue sociologique. Par exemple, la synthèse de Xia et Kirby⁶⁴

59. *Ibid.*, p.14.

60. Mieke Goos, Joana Pipa, Francisco Peixoto, « Effectiveness of grade retention: A systematic review and meta-analysis », art. cit.

61. C'est le cas dans celle réalisée par Karl L. Alexander, Doris R. Entwisle, Susan L. Dauber, *On the Success of Failure*, op. cit.

62. Andrew Abbott, *Time Matters: On Theory and Method*, Chicago, The University of Chicago Press, 2001.

63. Hugues Draelants et Sonia Revaz, *La politique des preuves en éducation*, op. cit.

64. Nailing Xia et Sheila Nataraj Kirby, *Retaining Students in Grade: A Literature Review of the Effects of Retention on Students' Academic and Nonacademic Outcomes*, Santa Monica, RAND Corporation, 2009.

montre les effets positifs du redoublement sur quelques années. Mais les auteurs concluent qu'« *une des plus importantes questions concernant les effets du redoublement est de savoir si les effets positifs à court terme persistent sur le long terme* »⁶⁵. La synthèse de Gary-Bobo et Robin est dans la même veine puisqu'après avoir fait état d'effets positifs du redoublement sur quelques années, les auteurs indiquent qu'ils ne seraient que transitoires et déconseillent dès lors son usage au regard de ses effets négatifs à *long terme*, notamment sur l'insertion socioprofessionnelle (plus difficile) et sur le salaire reçu (significativement inférieur à celui des élèves n'ayant jamais redoublé)⁶⁶.

Sur ce point encore, les critères de jugement utilisés par les chercheurs pour évaluer l'efficacité du redoublement tranchent avec ceux des acteurs scolaires qui attendent simplement du redoublement qu'il permette à un élève de se relancer temporairement : c'est l'idée d'un effet rebond qu'évoquent souvent les enseignants, les élèves et leurs parents. Force est de constater que les critères utilisés dans les études quantitatives ne sont pas neutres, mais implicitement porteurs d'une visée normative. On peut donc se demander s'il est sociologiquement pertinent d'exiger d'un redoublement qui intervient par exemple en fin de maternelle, en primaire ou en début de secondaire, qu'il produise encore des effets à long terme, jusqu'en fin de scolarité ou lors de l'entrée dans l'emploi. Cherche-t-on par le redoublement à changer un destin social entier ou à relancer un élève ? On peut certes souligner que le redoublement a des conséquences sur les carrières scolaires et touche les élèves inégalement selon leur origine sociale ou leur genre⁶⁷. Mais ce constat doit être replacé dans le fonctionnement du système éducatif. En Belgique, l'école répartit inégalement les biens et les positions scolaires, ce qui distingue socialement et professionnellement les élèves à leur sortie de l'école. Un chercheur qui déplore que le redoublement accentue les inégalités sociales doit se demander s'il n'est pas plutôt en train de questionner un modèle éducatif (c'est la fonction de sélection de l'École qui semble ici en cause, plus que le redoublement en tant que tel). En faisant cet amalgame, les études exigent beaucoup du redoublement : alors même qu'il est déjà complexe d'établir ses effets sur l'année scolaire, les critères

65. *Ibid.*, p.194.

66. Robert Gary-Bobo et Jean-Marc Robin, La question des redoublements. Analyse économique et problèmes statistiques, *Revue économique*, vol. 6, N°1, 2014, p.5-45.

67. Nailing Xia et Sheila Nataraj Kirby, *Retaining Students in Grade*, *op. cit.*

utilisés pour évaluer son efficacité suggèrent implicitement que le redoublement devrait encore faire effet parfois plus d'une dizaine d'années après l'événement.

Par ailleurs, il faut aborder la question de l'échelle temporelle. La méta-analyse de Holmes et Matthews utilise des études qui s'étalent de 1929 à 1981. Plus récemment, la méta-analyse de Goos, Pipa et Peixoto synthétise des études produites entre 2004 et 2021, mais intègre néanmoins des études portant sur des cohortes d'élèves scolarisées avant le *No Child Left Behind Act* de 2001⁶⁸. Il est nécessaire que les méta-analyses se fondent sur des contextes historiques cohérents. En l'occurrence, comparer des études ayant eu lieu dans un même système éducatif, mais en amont et en aval d'une réforme éducative majeure ne fait pas sens. En 30 ans, les changements ont été trop importants pour penser que les connaissances sont encore valables. Des études obsolètes ne devraient pas servir de référence pour asseoir un discours sur le redoublement. Cet argument est transposable aux sciences sociales. Aujourd'hui, personne ne dirigerait une politique de prévention sur le suicide à partir des travaux de Durkheim sans prendre le temps de les recontextualiser ou prendre connaissance d'études plus récentes.

En outre, et pour poursuivre sur la thématique de la temporalité, il est également important de questionner le fait que les méta-analyses agrègent à la fois des études fondées sur des suivis de cohortes pendant plusieurs années, parfois au-delà de la décennie⁶⁹, avec des travaux qui estiment l'effet du redoublement à partir d'événements singuliers comme des épreuves communes⁷⁰ ou des résultats PISA⁷¹. D'une part, aborder un événement tel que le redoublement hors de la séquentialité d'une scolarité présente des limites : « *les événements dans leur enchaînement ont des effets sur la construction des phénomènes* »⁷². Mais il y a une autre limite qui est celle d'agréger des approches processuelles et statiques des phénomènes et tient à deux conceptions radicalement opposées

68. Cette loi visait à améliorer la performance des élèves en fixant des normes nationales pour l'éducation et en exigeant des tests standardisés pour évaluer les progrès des élèves. Il s'agit donc d'un changement significatif dans la politique scolaire américaine.
69. Karl L. Alexander, Doris R. Entwisle et Susan L. Dauber, *On the Success of Failure*, op. cit.; Jan Hughes, Stephen West, Hanjoo Kim et Shelby Bauer, « Effect of early grade retention on school completion: A prospective study », *Journal of Educational Psychology*, vol. 110, n°7, 2018, p. 974-991.
70. Mohamed-Badrane Mahjoub, « The treatment effect of grade repetitions », *Education Economics*, vol. 25, n°4, 2017, p. 418-432.
71. Michèle Belot et Vincent Vandenberghe, « Evaluating the 'threat' effects of grade repetition: exploiting the 2001 reform by the French-Speaking Community of Belgium », *Education Economics*, vol. 22, n°1, 2014, p. 73-89.
72. Andrew Abbott, *Time Matters*, op. cit., p. 9.

du redoublement. La question qui se pose est ici celle de l'horizon temporel des variables, et plutôt que de chercher à comprendre les phénomènes de causalité, il s'agit alors de se demander dans quelle mesure ou à quel moment les variables cessent de faire effet⁷³. C'est ici tout l'enjeu de l'encodage d'événements passés dans les phénomènes étudiés présentement. Par exemple, dans le cas des attitudes scolaires, celles-ci ont pu causer le redoublement, et s'être aggravées par la suite : un manque d'attention ou de participation en classe a pu apparaître avant le redoublement. C'est ce qu'Abbott appelle l'« *ambiguïté durationnelle* »⁷⁴ : l'indicateur observé a une durée inconnue dans le temps, ce qu'une étude ponctuelle peut difficilement étudier.

Un troisième critère sociologiquement discutable des études synthétisées par les méta-analyses tient à la manière dont on considère la promotion automatique à laquelle le redoublement est toujours comparé. La méthodologie des études sur le redoublement est en effet basée sur un raisonnement contrefactuel qui cherche à évaluer ce qui se serait passé si les élèves étaient traités différemment, s'ils avaient été promus plutôt que d'avoir redoublé. Mais en réalité, seul le redoublement est considéré comme un « traitement ». La promotion automatique en revanche est implicitement perçue comme un dispositif neutre. Pourtant, il n'y a rien d'anodin à être promu d'une année à l'autre tout en étant en échec scolaire. La neutralité accordée par certains à la promotion s'appuie sur l'idée que la norme scolaire serait de progresser dans la classe supérieure année après année. C'est effectivement le cas pour l'élève qui a acquis les apprentissages au programme et qui est donc en réussite scolaire. En revanche, passer dans l'année supérieure en cas d'échec n'est pas neutre du point de vue des acteurs scolaires. Cela dépend en fait des contextes. Dans certains systèmes, comme dans celui de la Belgique francophone, la norme est plutôt d'être maintenu, lorsqu'on n'a pas acquis les apprentissages requis. Dans un tel contexte, promouvoir un élève en échec revient à mettre en cause l'évaluation scolaire et le contrat sur lequel fonctionne habituellement l'école. Autrement dit, les chercheurs qui présumement cette neutralité font fi, une fois encore, des raisonnements pratiques des personnels éducatifs ou, pour reprendre les termes de Dodier et Barbot, du « travail normatif » que les professionnels

73. *Ibid.*

74. *Ibid.*

réalisent autour du dispositif du redoublement⁷⁵. Ils ne décrivent pas l'école telle qu'elle est, mais telle qu'ils voudraient qu'elle soit.

Un quatrième critère sociologiquement discutable est celui de la prévalence de la note dans les conditions d'appariement ou d'appariement entre élèves. Une telle décision revient en fait à placer les acteurs scolaires dans un rôle secondaire dans la construction des analyses du redoublement, puisque les analyses se fondent sur les points obtenus par les élèves. Cela peut venir du fait que l'on conçoit les acteurs comme «*des automates rationnels*» selon Abbott⁷⁶. Ce qui nous semble d'ailleurs le cas dans la mesure où les chercheurs supposent par exemple que les acteurs, si on leur présente des connaissances scientifiques, vont changer de pratiques⁷⁷. Dans ce cas de figure, le motif de délibération des enseignants n'est pas pris en compte, alors même que l'échec scolaire n'est pas la seule raison d'attribution du redoublement, et intervient même parfois de manière secondaire face à des évaluations comportementales⁷⁸. Ce faisant, une partie des événements qui interviennent à l'origine du phénomène étudié est laissée de côté dans les approches classiques.

En résumé, d'un point de vue sociologique, on est fondé à s'interroger sur les choix qui ont présidé aux critères des études quantitatives synthétisées par les méta-analyses. Les enseignants, les équipes pédagogiques n'évaluent pas le redoublement de la même manière que les chercheurs. Les choix des uns et des autres répondent à des rationalités différentes, qui possèdent chacune leur logique propre, mais ils ne sont pas neutres. Or les choix des chercheurs sont habituellement présentés et perçus comme purement scientifiques, ce qui est trompeur comme nous l'avons vu.

Pour conclure ce point et éviter les malentendus, précisons que ce qui nous paraît problématique dans les approches positivistes du redoublement n'est pas la formalisation. Rappelons qu'observer le réel, c'est toujours abstraire, identifier certaines proprié-

75. La manière dont les individus évaluent un dispositif ou n'importe quel état de choses sont désignées par Dodier et Barbot comme étant du travail normatif au sens où « toute évaluation se réfère en effet nécessairement, quoique plus ou moins explicitement, à des attentes normatives concernant ces états de choses » (p. 426). Voir Nicolas Dodier et Janine Barbot, « La force des dispositifs », *Annales - Histoire, Sciences Sociales*, vol. 71, N°2, 2016, p. 421-448.

76. Andrew Abbott, *Time Matters*, op. cit., p. 16.

77. Marcel Crahay, *Peut-on lutter contre l'échec scolaire?* Bruxelles, De Boeck, 2003.

78. Géry Marcoux et Marcel Crahay, « Mais pourquoi continuent-ils à faire redoubler ? Essai de compréhension du jugement des enseignants concernant le redoublement », *Schweizerische Zeitschrift für Bildungswissenschaften*, vol. 30, N°3, 2008, p. 501-518 ; Caroline De Pascale et Hugues Draelants, « Le redoublement est une décision d'équipe : comprendre le sens de l'agir enseignant face au redoublement », *Éducation et Sociétés*, vol. 50, N°2, 2023, p. 161-178.

tés et en ignorer d'autres. En sciences, l'abstraction est donc inévitable⁷⁹. Mais elle peut être plus ou moins pertinente. Les propriétés qu'il convient de privilégier sont celles qui, a posteriori, se révèlent utiles pour mettre de l'ordre dans les faits et les comprendre. C'est précisément là que le bât blesse. Dans le cas présent, les études se fondent sur de mauvaises abstractions ou en tout cas des abstractions sociologiquement questionnables dans la mesure où elles impliquent l'usage de concepts ou de critères de jugement en décalage avec le travail normatif des professionnels. Cette perspective sociologique, mais aussi pragmatique, au sens où elle se soucie de la mise en œuvre de l'action publique va être clarifiée dans la partie qui suit consacrée à la réception des politiques de lutte contre le redoublement. Politiques qui justement ont échoué jusqu'à présent, faute de compréhension sociologique et de prise en compte de la légitimité pragmatique, puisque l'expertise qui sous-tend les réformes s'inspire fortement d'approches positivistes.

Les MA sont-elles une bonne source de légitimité pour l'action publique?

Le cas de l'échec des politiques de lutte contre le redoublement en Belgique francophone

La question de savoir si les réformes politiques fondées sur les preuves dont nous venons de discuter atteignent leurs objectifs est cruciale, puisqu'elles sont utilisées non seulement pour guider l'action publique, mais aussi pour la légitimer, et transformer les pratiques notamment des enseignants. Et il faut partir ici du constat que trente années de politiques visant à limiter le redoublement ne sont pas parvenues à le réduire significativement.

La Belgique compte en effet l'un des plus grands nombres d'enfants ayant redoublé dans le monde occidental : 30,2 % des élèves de 15 ans ont redoublé dans l'ensemble de la Belgique. Dans le même temps, 46 % des jeunes de 15 ans de la Fédération Wallonie-Bruxelles (ou FWB, la partie francophone de la Belgique) ont redoublé. La FWB a le taux de redoublement le plus élevé de l'OCDE. À titre de comparaison, la moyenne de l'OCDE est de 11 %⁸⁰. Le

79. Voir l'excellent dossier en deux volumes que *L'Année sociologique* a consacré à cette question et notamment l'article introductif de Pierre Demeulenaere, « Présentation : Le problème de l'abstraction en sociologie », *L'Année sociologique*, vol. 56, n°2, 2006, p. 253-262.

80. Programme for International Student Assessment, OCDE, 2020.

redoublement a généré un coût supplémentaire d'environ 367 millions d'euros pour l'année 2022⁸¹. Ces taux de redoublement hors-norme sont donc devenus une préoccupation centrale des politiques éducatives en Belgique depuis le début des années 1990. À ce jour, pas moins de trois textes législatifs différents ont explicitement visé à réduire ces taux de redoublement dans l'enseignement secondaire⁸² (voir tableau 2).

Nom du texte	Contenu
Arrêté du Gouvernement de la Communauté française du 19 juillet 1993 modifiant l'arrêté royal du 29 juin 1984 relatif à l'organisation de l'enseignement secondaire	Passage automatique de la première à la deuxième année de l'enseignement secondaire, suppression du redoublement pour les deux premières années de l'enseignement secondaire. À partir de la rentrée 1995-1996, le premier diplôme doit être obtenu en 3 ans maximum, avec la possibilité d'une année supplémentaire à la fin de la deuxième année qui est censée être différente du simple redoublement.
Décret relatif à l'organisation du 1er degré de l'enseignement secondaire du 19 juillet 2001 (M.B. du 23 août 2001).	Abandon du principe de promotion automatique, introduction d'une année complémentaire qui « ne peut en aucun cas constituer une répétition de l'année précédente » (p. 1). À partir de la rentrée 2001-2002, un élève peut compléter le 1er degré en 3 ans maximum et a la possibilité de faire des années complémentaires à la fin de la première année (1AC/1S) ou après la deuxième année (2CC/2S).
Circulaire n°5795 du 30 juin 2016 Circulaire générale relative à l'Organisation de l'enseignement secondaire ordinaire et à la Sanction des études.	Suppression de la 1re année complémentaire dans l'enseignement secondaire à partir de la rentrée scolaire 2015-2016. Les élèves peuvent encore effectuer une deuxième année complémentaire (2S) à la fin de leur deuxième année d'enseignement secondaire.

Tableau 2. Législations visant à réduire le redoublement dans l'enseignement secondaire en FWB depuis 1990

Avant de revenir sur ce qui a été ignoré par l'action publique en se fondant sur les méta-analyses, il nous faut expliquer comment concrètement elles ont été mobilisées pour fonder la politique de lutte contre le redoublement. Précisons que le lien est toujours médié : jamais les débats parlementaires n'évoquent en tant que telles les méta-analyses. Les responsables politiques évoquent plu-

81. Fédération Wallonie-Bruxelles, *Les indicateurs de l'enseignement 2023*, Bruxelles, Direction générale du Pilotage du Système éducatif, 2023.
82. Un schéma similaire peut être observé dans l'enseignement primaire. Cependant, nous limitons ici notre propos aux trois premières années de l'enseignement secondaire, car c'est là que l'on trouve les taux de redoublement les plus élevés. En outre, c'est surtout sur les deux premières années de l'enseignement secondaire que se sont concentrées les tentatives de réduction du redoublement. Nos propres recherches empiriques ont donc porté sur ce segment de la scolarité.

tôt « la recherche » ou les rapports que le politique a commandés pour s'informer de l'état de la recherche sur le sujet et le nom des experts mandatés, auteurs de ces rapports. Notre analyse des débats parlementaires et des trois rapports d'expertises commandités en trente ans sur la question du redoublement et présentés aux députés montre la place prépondérante de ces méthodes, ainsi que des comparaisons internationales.

Le premier rapport⁸³, commandité en 1993 se fonde sur la comparaison internationale (la Belgique tient le record des redoublements d'Europe), et mentionne le coût de ce dispositif. Il évoque ensuite particulièrement les recherches sur le redoublement : la synthèse narrative de Jackson, et la méta-analyse de Holmes et Matthews. Deux autres études individuelles sont également citées, l'une menée en France, l'autre réalisée dans le canton de Genève. En revanche, aucune recherche conduite en Belgique sur le redoublement n'est citée. Ce rapport qui aboutit à une condamnation sans appel du redoublement et qui sera utilisé pour légitimer la lutte contre le redoublement en Belgique francophone se fonde donc uniquement sur des études menées dans d'autres contextes. Un deuxième rapport⁸⁴, commandité en 2011, sur le cas du redoublement dans l'enseignement primaire donne, lui aussi, une place prépondérante aux méta-analyses et notamment à celle de Jimeron⁸⁵ pour souligner que le redoublement a des effets négatifs et qu'il faudrait l'abandonner. Il en va de même dans le troisième rapport d'expertise⁸⁶ qui, bien qu'il mobilise une bibliographie plus étoffée que le rapport de 1993, continue d'une part d'invoquer la comparaison internationale et d'autre part de se fier essentiellement aux études quantitatives et aux méta-analyses (en citant encore une fois les trois plus anciennes) pour tirer des conclusions sur le phénomène du redoublement. Cette ignorance du contexte, qui postule implicitement une validité universelle des résultats obtenus dans ces études⁸⁷, est pour le moins discutable.

83. Marcel Crahay, « Échec des élèves, échec de l'école? », *Recherche en Éducation. Théorie et Pratique*, vol. 11-12, 1993.

84. Sabine Kahn, Bernard Rey, Charlotte Bouko et Sylvie Vanlint, *Analyse des causes et conséquences du maintien en troisième maternelle en Fédération Wallonie-Bruxelles*, Rapport pour le Service général du Pilotage du système éducatif, 2011.

85. Shane Jimeron, « Meta-analysis of Grade Retention Research: Implications for Practice in the 21st Century », art. cit.

86. Ariane Baye, Florent Chenu, Marcel Crahay, Dominique Lafontaine et Christian Monseur, *Le redoublement en Fédération Wallonie-Bruxelles*, Rapport d'expertise commandité par le groupe parlementaire du Parti Socialiste, 2014.

87. Amy L. Reschly et Sandra L. Christenson, « Grade retention: Historical perspectives and new research », *Journal of school psychology*, vol. 51, n°3, 2013, p. 319-322.

Après ces trois réformes, quelles conclusions les autorités scolaires de la FWB ont-elles tiré ? En 2007, la FWB notait « *Le faible impact de cette réforme [de 1993] sur la fin des études secondaires implique une production simultanée et plus importante de redoublement en secondaire 4, 5 et 6* »⁸⁸. Ces constats d'échec de la réforme ont conduit la députée (et future ministre de l'Éducation) Caroline Désir à remettre un rapport au parlement en 2014, qui soulignait : « *Aujourd'hui, les taux de redoublement sont plus élevés qu'en 1992-1993* ». Enfin, quelques années plus tard, notons l'absence d'effets depuis la suppression de la 1^{re} année complémentaire en 2015, car « *il y a une nette augmentation de la courbe du taux de redoublement pour les élèves qui font une deuxième année* »⁸⁹.

Nous avons analysé ailleurs en détail les raisons pour lesquelles les enseignants continuent d'adhérer au redoublement, malgré le discours politique et académique décrivant ses effets délétères sur les élèves. Cette analyse a montré que le redoublement remplissait diverses « *fonctions latentes* »⁹⁰. Ainsi, il sert notamment à gérer l'hétérogénéité des élèves en les triant au sein des établissements. Il joue par ailleurs un rôle dans la régulation de l'ordre scolaire et du comportement des élèves. Les écoles l'utilisent aussi pour se distinguer des écoles voisines. Enfin, il contribue au maintien de l'autonomie professionnelle des enseignants en leur permettant de décider du sort des élèves en fin d'année. Plus récemment, nous avons analysé les modalités d'attribution du redoublement par les enseignants, et les fonctions manifestes qu'il servait. Il ressort que, outre l'avis de l'enseignant sur le redoublement, c'est le contexte du conseil de classe et de la délibération collective des cas d'élèves dans des situations similaires qui détermine les pratiques enseignantes⁹¹. En outre, nous avons constaté une grande hétérogénéité des causes et objectifs du phénomène selon les contextes scolaires : les enseignants vont tantôt viser la réussite du curriculum et des études supérieures, tantôt viser une réorientation en début de cycle⁹², une telle hétérogénéité rend difficilement comparables des redoublements.

88. « Indicateurs de l'enseignement », Ministère de la Communauté Française de Belgique, 2007.

89. « Indicateurs de l'enseignement », Ministère de la Communauté Française de Belgique, 2020, p. 36.

90. Hugues Draelants, « Les fonctions latentes du redoublement », *Éducation et Sociétés*, n°21, 2008, p. 163-180.

91. Caroline De Pascale, Hugues Draelants, « Le redoublement est une décision d'équipe : comprendre le sens de l'agir enseignant face au redoublement », art. cit.

92. *Ibid.*

À partir de ces travaux et des constats d'échec des politiques éducatives portant sur le redoublement ainsi que de leur place dans la construction du redoublement comme problème, il nous semble important de creuser les raisons pour lesquelles cette méthode limite et entrave l'action publique en éducation.

Des difficultés politiques issues de l'usage des méta-analyses

Nous associons l'échec des politiques éducatives relatives au redoublement en Belgique francophone à un manque de connaissance des usages locaux de ce dispositif, faute d'études portant sur ce contexte reprises dans les rapports d'expertise et qui auraient permis de comprendre les fonctions latentes que remplit le redoublement et d'anticiper la difficile réception de ces politiques sur le terrain. Il s'agit ici d'un inconvénient lié à ces abstractions sociologiquement questionnables dont nous parlions plus haut, puisque les recherches sur le redoublement l'évaluent principalement à l'aune de critères distincts de ceux utilisés par les enseignants pour décider d'un redoublement puis pour l'évaluer. Ce faisant les causes pragmatiques de l'usage du redoublement ne sont pas traitées par la réforme politique. Ceci occasionne un défaut de connaissance sur le phénomène, dont l'origine est liée entre autres à la théorie du changement qui a guidé l'action publique⁹³. Celle-ci a fait le pari que la démonstration scientifique de l'inefficacité du redoublement, conclusion à laquelle parviennent les méta-analyses sur le phénomène, suffirait à convaincre les enseignants du bien-fondé de la réforme⁹⁴. Or, la sociologie de l'action publique nous rappelle que l'implémentation des politiques est un enjeu clé de l'action publique, celui qui en détermine le sens en fin de compte puisque, sans considérer la mise en œuvre, « *on ne saisit au mieux que des annonces politiques* »⁹⁵.

Plusieurs types de légitimité doivent être atteints pour faciliter la mise en œuvre d'une réforme : la légitimité morale qui « *renvoie à une adéquation entre les nouvelles idées portées par le changement et les schémas normatifs partagés et dominants* »⁹⁶ ; la légitimité cognitive qui « *fait plutôt écho à une adéquation entre les idées*

93. Hugues Draelants, *Réforme pédagogique et légitimation. Le cas d'une politique de lutte contre le redoublement*, Bruxelles, De Boeck, 2009.

94. *Ibid.*

95. Pierre Lascombes et Patrick Le Galès, *Sociologie de l'action publique*, Paris, Armand Colin, 2018.

96. Hugues Draelants et Sonia Revaz, *Penser la plausibilité des réformes éducatives. La légitimation comme condition du changement institutionnel*, Rapport pour le CNESECO, 2022, p. 22.

*promues par le changement et les connaissances (notamment issues de la recherche)»⁹⁷ et enfin la légitimité pragmatique ou légitimité de terrain qui renvoie « à une démonstration de la supériorité opérationnelle des nouveaux modes de fonctionnement »⁹⁸. En l'occurrence, les discours politiques fondés sur la parole des experts et la condamnation large dont le redoublement fait l'objet leur permettent d'asseoir la légitimité cognitive et la légitimité morale de la lutte contre le redoublement, mais pas sa légitimité pragmatique. Or, comme nous l'avons écrit ailleurs : « *S'il est évidemment indispensable que les professionnels de l'école soient convaincus du bien-fondé et de la pertinence d'une réforme, nous estimons qu'il importe encore davantage qu'ils puissent constater de son efficacité et des bénéfiques qu'elle leur permet de tirer. En effet, ce sont bien les pratiques que les réformes ambitionnent de réguler et pas juste les représentations ou les croyances des individus.* »⁹⁹*

Grâce à son statut privilégié pour les décideurs politiques, au sommet de la hiérarchie de la preuve, une méta-analyse donne un surcroît de légitimité cognitive, voire morale, aux annonces de réformes politiques ce qui à l'échelle temporelle d'un mandat fonctionne, mais l'absence de légitimité pragmatique des réformes complique considérablement l'implémentation. C'est précisément parce qu'elles se fondent sur des études qui, comme nous l'avons vu, sont en décalage avec les critères d'évaluation qui prévalent sur le terrain pour juger de la pertinence et de l'efficacité d'un redoublement que les connaissances fournies par les méta-analyses ne procurent pas de légitimité pragmatique à l'action publique. Pour tenter d'établir les conditions de légitimité pragmatique d'une politique visant à réduire le redoublement, d'autres connaissances, plus qualitatives et plus ancrées dans les savoirs professionnels nous paraissent nécessaires.

Puisque, sur le terrain, les professionnels distinguent le « bon » du « mauvais » redoublement, sans doute serait-il pertinent de tirer parti de leur expertise contextuelle pour proposer des changements certes plus modestes, mais plus praticables, car plus compréhensibles et finalement probablement plus efficaces s'ils sont mieux acceptés des acteurs de terrain. Cela suppose de prendre une certaine distance avec le modèle de l'éducation par les preuves, qui monte actuellement en puissance et porte au pinacle

97. *Ibid.*, p. 22.

98. *Ibid.*, p. 22.

99. *Ibid.*, p. 7.

la méthode de la méta-analyse, sans faire droit à la pluralité des savoirs et des approches des recherches en éducation ni aux pratiques et savoirs des professionnels¹⁰⁰. Il importe donc d'évaluer l'effet du redoublement en tenant compte aussi de ces derniers que des recherches sociologiques et plus qualitatives permettent de cerner, surtout si l'enjeu est d'appeler à une diminution/suppression de son emploi. Étant donné que le redoublement remplit d'autres fonctions que pédagogique¹⁰¹, il est vain de vouloir changer les usages du redoublement simplement en modifiant les représentations des enseignants, par exemple lors de leur formation initiale, sans tenir compte des conditions effectives de travail et des contraintes qu'elles induisent sur leurs représentations¹⁰².

Cependant, une des raisons pour lesquelles les savoirs professionnels ne sont pas non plus inclus dans les approches quantitatives et notamment dans les méta-analyses – dans lesquelles il n'y a pas de catégorie liée au jugement enseignant –, c'est aussi parce que les études portant sur la question et ensuite synthétisées abordent les pratiques enseignantes sous l'angle des perceptions¹⁰³, des croyances¹⁰⁴ ou des points de vue¹⁰⁵. Plus rarement, les auteurs entendent comprendre ce qui se fait sur le terrain¹⁰⁶. Un tel vocabulaire n'est pas anodin puisqu'il renforce l'opposition croyances/connaissances, en entretenant le clivage entre les chercheurs qui inspirent l'action publique, analysent les phénomènes

100. *Ibid.*

101. Hugues Draelants, « Les fonctions latentes du redoublement », art. cit.

102. Sylvain Broccolichi, Christophe Joigneaux et Catherine Couturier, « Comment les jeunes maîtres appréhendent-ils les difficultés de leurs élèves ? Aspirations initiales, obstacles actuels et possibilités de les surmonter », *Recherche & formation*, vol. 87, n°1, 2018, p. 29-45.

103. Amira Mattison, Linda Raffaele Mendez, Robert Dedrick, Sarah Dickinson, Emily Wingate et Camille Hanks, « Early elementary teacher ratings of behavior as predictors of grade retention: Race, gender, and socioeconomic status as potential moderators », *Psychology in the Schools*, vol. 55, n°10, 2018, p. 1171-1187.

104. Voir Marcel Crahay, Élizabéth Issaieva et Christian Monseur, « Les enseignants face au redoublement : ceux qui y croient et ceux qui n'y croient pas. Étude de leurs conceptions psychopédagogiques et de leur connaissance des recherches », *Revue Française de Pédagogie*, vol. 187, n°2, 2014, p. 35-53 ; Elizabeth Walton, « Teacher beliefs about grade repetition: An exploratory South African study », *Citizenship Teaching and Learning*, vol. 13, n°1, 2018, p. 45-60 ; Natalie N. Santos et Vera Monteiro, « Second grade retention: Beliefs, decision-making styles, and factors involved in the decision process », *Psychology in the Schools*, vol. 61, n°4, 2024, p. 1646-1666.

105. Suzanne Young, Natacha Trujillo, Marie Alice Bruce, Tracie Pollard, Jeff Jones et Bret Range, « Preservice teachers' views About grade retention as an intervention for struggling students », *Preventing School Failure*, vol. 63, n°2, 2019, p. 113-120.

106. Géry Marcoux et Marcel Crahay, « Mais pourquoi continuent-ils à faire redoubler ? Essai de compréhension du jugement des enseignants concernant le redoublement », art. cit. ; Daniel Owusu Kyereko, William C. Smith, Ishmael Hlovor et Gabriel Keney, « Understanding grade repetition from the perspectives of teachers and principals in basic schools in Ghana », *International Journal of Educational Development*, vol. 93, 2022, p. 1-9.

selon certains critères légitimes, et les enseignants perçus comme en défaut de connaissance, de rationalité, et supposés n'être que des exécutants¹⁰⁷.

Une autre raison de l'échec des politiques éducatives portant sur le redoublement tient au recours à la science comme manière de dépolitiser et de confisquer le débat. Il est temps de souligner que la suppression du redoublement est en ce sens un changement dont on sous-estime l'ampleur. Les présuppositions sur lesquelles se fondent certains auteurs, et qui se traduisent (on l'a vu) notamment par le fait d'envisager la promotion automatique non pas comme un traitement spécifique, mais comme une action neutre, sont trompeuses. Il est faux de penser que le redoublement serait une pratique dont la suppression n'aurait que des conséquences pour les élèves concernés. Les conséquences (positives ou négatives) d'une éventuelle suppression du redoublement, là où il est traditionnellement utilisé, concernent l'ensemble des élèves et le système éducatif tout entier¹⁰⁸.

On ne souligne pas assez combien la suppression du redoublement équivaut à un changement de paradigme ou de philosophie de l'éducation, que l'on peut résumer comme celui du renoncement à la fonction de sélection de l'école ou à la mise entre parenthèses de la méritocratie dans l'école commune, ce qui est un changement profond et qui s'opère sans réel débat politique. Le redoublement n'est pas juste un dispositif de remédiation (à l'efficacité controversée). C'est la garantie que tout travail (scolaire) mérite d'être rémunéré, et sera récompensé par une réussite ou par un passage dans la classe supérieure¹⁰⁹. Parallèlement, sa suppression revient à dissocier la réussite et le travail scolaire : l'école sans redoublement n'est plus une école où l'élève doit réussir pour avancer, ce qui impliquait généralement de travailler, de faire la preuve que l'on avait fourni des efforts pour répondre aux attentes scolaires. Ce sont donc aussi les attentes scolaires qui se voient complètement modifiées. Cela n'échappe pas aux professionnels qui s'en plaignent volontiers¹¹⁰.

107. Christian Maroy et Branka Cattonar, « Professionnalisation ou déprofessionnalisation des enseignants ? Le cas de la Communauté française de Belgique », *Les Cahiers de recherche du GIRSEF*, n°18, 2002.

108. Hugues Draelants et Jacques B. Georges, « La fonction morale du redoublement », *Éducation et Sociétés*, à paraître.

109. *Ibid.*

110. Hugues Draelants, *Réforme pédagogique et légitimation*, *op. cit.*

Au-delà de l'importance d'inscrire cette discussion des finalités dans les politiques nationales ou dans les professionnalités locales, il y a une responsabilité de la communauté scientifique internationale concernant l'évolution des finalités de l'éducation. Les liens entre études internationales, idéologies et politiques éducatives existent, mais sont encore complexes à comprendre¹¹¹. Ces liens, et l'absence d'autres formes de gouvernance internationale en matière d'éducation donnent davantage de poids aux décisions des chercheurs qui publient des études visant à promouvoir des « *best practices* » à l'échelle globale. Il y a un contexte actuel de Nouvelle Gestion Publique marqué par la montée en puissance d'une volonté de standardiser les pratiques enseignantes jugées « efficaces » sur critères d'études internationales¹¹². La montée d'un État régulateur, voire évaluateur qui se décline selon les modèles nationaux,¹¹³ renforce la propension des États à utiliser des indicateurs quantitatifs pour évaluer les établissements, les acteurs, mais aussi les acquis des élèves, c'est-à-dire les « produits » du système éducatif. Dans un tel contexte, étudier le redoublement à partir de critères peu (ou pas) discutés, d'un cadre théorique qui a peu (ou pas) de connaissance des pratiques professionnelles, d'un avis sur le redoublement fondé sur des travaux obsolètes, tient presque d'une pratique frauduleuse de la science¹¹⁴. Et il faudrait plutôt admettre que les preuves sont actuellement insuffisantes pour informer correctement les décideurs au sujet du redoublement¹¹⁵.

Enfin, il faut critiquer la logique uniforme avec laquelle les politiques portant sur le redoublement sont conçues, il s'agit toujours en pratique de le limiter ou de le supprimer, au profit du passage dans l'année supérieure, sans que l'entièreté des fonds économisés par cette mesure ne soit réinjectée dans des moyens de remédiation. Vu comme une mesure néfaste, il est séduisant de produire un discours de rationalisation budgétaire qui supprime le redoublement et, *de facto* ce qu'il coûte à la société. Les

111. Manuel Enrique Cardoso, « Policy evidence by design: International large-scale assessments and grade repetition », *Comparative Education Review*, vol. 64, n°4, 2020, p. 598-618.
112. Branka Cattanar et Vincent Dupriez, « Recomposition des professionnalités et de la division du travail enseignant en situation d'obligation de rendre des comptes. Le cas des professionnels de l'éducation en Belgique francophone », *Éducation et Sociétés*, vol. 43, n°1, 2019, p. 25-39.
113. Christian Maroy, « Vers une régulation post-bureaucratique des systèmes d'enseignement en Europe ? » *Sociologie et sociétés*, vol. 40, n°1, 2008, p. 31-55.
114. John P.A. Ioannidis, « The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses », art. cit.
115. Hugues Draelants, « Le redoublement est-il vraiment moins efficace que la promotion automatique ? », art. cit.

méta-analyses se gardent pourtant bien de préconiser l'abandon pur et simple du redoublement et insistent sur la place à accorder à l'individualisation des apprentissages et à la remédiation scolaire, c'est notamment le cas de celle de Jimerson qui est la plus citée. Cependant, on ne peut nier l'effet du *design* par lequel le redoublement est maintenant abordé depuis quarante ans dans les méta-analyses : celle où on oppose le redoublement à la promotion automatique, supposée avoir des effets neutres sur la carrière scolaire. Une logique du cas par cas semble préférable pour tenir compte des singularités et spécificités contextuelles.

Pour aller plus loin, nous pensons que l'échec des politiques éducatives en matière de redoublement doit constituer un point de départ pour explorer les conditions qui amènent un élève, plutôt qu'un autre, à bénéficier du redoublement. Il y a un enjeu à retourner au niveau local et aux diverses réalités qu'y s'y trouvent, mais aussi à comprendre que le contexte ne se réduit pas à des échelles nationales. Le but serait aussi d'approfondir la connaissance des facteurs susceptibles d'influencer les résultats du redoublement. En somme, nous pensons qu'il faut élargir la notion de preuve sur laquelle se fonde souvent l'action publique. *Evidence based education* ne doit pas déprofessionnaliser les enseignants ou les déposséder des savoirs valorisés. Or, c'est précisément l'apport des méthodes mixtes ou qualitatives que de permettre de comprendre ces contextes locaux à l'origine des résultats non concluants auxquels les méta-analyses font face. Enfin, il est important de dissocier la mesure d'un phénomène de l'évaluation de celui-ci. La mesure n'est pas anodine, pas plus que les critères sur lesquels elle se fonde.

Conclusion

Une critique nécessaire de la hiérarchie de la preuve

Les méta-analyses ont initialement pour ambition d'aborder l'échelle des systèmes éducatifs. Cependant, l'aspect multidimensionnel et mouvant des finalités des systèmes éducatifs qui résultent des contestations régulières des politiques antérieures est mieux compris par des explications multi-théoriques¹¹⁶ et une triangulation entre les sources de données¹¹⁷. L'avantage de l'ana-

116. Stephen J. Ball, « What Is Policy? Texts, Trajectories and Toolboxes », art. cit.

117. Jo Helgetun, *Maintaining and changing institutionalized environments: Teacher education policymaking in England and France compared*, Doctoral Dissertation, UCLouvain, 2021.

lyse multi-source dans l'éducation est que diverses approches qui éclairent différents aspects des processus d'apprentissage, et les effets secondaires (de non-réalisation) d'une approche de l'apprentissage peuvent être complémentaires. Par conséquent, nous pensons qu'il est nécessaire d'adopter une approche multi-méthodologique pour analyser toutes les facettes du redoublement plutôt qu'une approche méthodologique unique ancrée dans quelque chose qui court le risque d'être dépendant de la qualité des études incluses et peut conduire à la diffusion de conclusions erronées et potentiellement problématiques pour l'action publique, puisque ces preuves se présentent comme directement utiles pour agir. Des limitations inhérentes aux méta-analyses ne peuvent être surmontées par ce type de méthodologies. En conséquence, réduire l'analyse du redoublement à une méta-analyse risque à la fois d'aveugler la communauté des chercheurs et de la priver d'informations inestimables qui ne sont pas prises en compte par cette méthode, mais aussi, ce qui est pire, de nous conduire à fonder nos recommandations sur une science sérieusement défectueuse. Ce qui ne serait en définitive pas une bonne manière de légitimer l'action publique.

D'après Biesta, l'éducation lie intrinsèquement les notions de valeur et d'efficacité. Ce qui marche le mieux dépend du but d'un système éducatif¹¹⁸. Ou, pour le dire en des termes plus durkheimiens, cela revient à considérer qu'il faut être au clair sur l'idéal moral d'une société pour identifier les défauts de régulation et d'intégration dans cette société. Le redoublement n'échappe pas à ce constat. Par exemple, le *English Education Endowment Foundation*¹¹⁹ a publié un rapport très négatif vis-à-vis du redoublement en septembre 2021, le considérant coûteux (le système éducatif anglais valorisant la notion de rentabilité), peu évalué positivement par les études (le rapport notait que la plupart des études sont menées aux États-Unis), et accentuait les inégalités scolaires. Pour le dire autrement, ce rapport s'est fondé sur des considérations pécuniaires, scientifiques et morales pour fonder ses conclusions. En Norvège les logiques sous-tendant l'éducation obligatoire (de 6 à 16 ans) sont celles d'homogénéité de l'offre, d'égalité d'accès et d'équité de participation, ce qui influence autant les politiques

118. Gert Biesta, « Good education in an age of measurement: on the need to reconnect with the question of purpose in education », *Educational Assessment, Evaluation and Accountability*, vol. 21, n°1, 2009, p. 33-46.

119. Voir : educationendowmentfoundation.org.uk/education-evidence/teaching-learning-toolkit/repeating-a-year, consulté le 27 mai 2022.

que les recherches éducatives¹²⁰. Ce sont d'ailleurs ces principes moraux qui sont avancés pour écarter toute discussion en faveur du redoublement. Une forme de redoublement (fondé sur les résultats aux épreuves) et de réorientation vers des filières d'enseignement apparaît cependant après l'âge de 16 ans. Le cas est similaire en Finlande où, après une éducation égalitaire jusqu'à l'âge de 15 ans, l'entrée dans l'enseignement universitaire peut s'avérer difficile pour les élèves les moins performants. Ce que nous pensons ici c'est qu'on ne peut superposer des logiques d'un État (par exemple la méritocratie française) sur celles d'un autre (telle que la logique d'équité sociétale nordique).

En outre, nous pensons qu'il serait judicieux de distinguer les critères rationnels avec lesquels on évalue le redoublement, des critères raisonnables, compte tenu des systèmes éducatifs, de leurs missions et de leurs moyens. C'est d'ailleurs pourquoi il faut également faire preuve de prudence avec la notion de preuve en éducation. Ce qui est considéré comme avantageux/nuisible pour les élèves dépend des buts des systèmes éducatifs et demande d'interroger les finalités des systèmes éducatifs.

Un succès injustifié

Pour toutes les raisons évoquées, la politique sur le redoublement souffre d'un déficit de légitimité qui se traduit par exemple par un diagnostic d'échec de la politique menée en Belgique francophone, ou en France (par le revirement récent sur le redoublement). Et si ce déficit de légitimité advient, c'est en grande partie du fait de la mauvaise abstraction évoquée précédemment. Donc, les promesses sur lesquelles repose le succès des approches *evidence-based* sont injustifiées. Elles procurent surtout un faux sentiment de certitude. Elles admettent comme critique la question de leur mise en œuvre, et un certain nombre d'études montrent que le passage des expériences à une plus large échelle ne donne pas forcément les résultats espérés et ils pensent qu'en travaillant à améliorer la mise en œuvre, conçue comme une science en soi,

120. Jo Helgetun, Justin J.W. Powell et Mike Zapp, «Governing Educational Research in Norway: coercion, cooperation or consensus?», in M. Zapp, M. Marques et J. Powell (eds.), *European Educational Research (Re)Constructed: institutional change in Germany, the United Kingdom, Norway, and the European Union*, Symposium Books, 2018, p. 91-111; Mike Zapp, Jo Helgetun et Justin J.W. Powell, «(Re)shaping educational research through "programmification": Institutional expansion, change, and translation in Norway», *European Journal of Education*, vol. 53, N°2, 2018, p. 202-217; Kirsten Sivesind et Berit Karseth, «An officially endorsed national curriculum: institutional boundaries and ideological concerns», *Curriculum Perspectives*, vol. 39, N°2, 2019, p. 193-197.

on peut y remédier¹²¹. Mais, au fond, notre analyse suggère que le problème est bien plus profond. Il ne se limite pas à la question de l'implémentation, c'est la question de la conception même de l'action publique qui pose problème et la nature de l'expertise mobilisée lors de cette étape. Dans un domaine comme l'éducation, les connaissances valorisées et produites par les recherches qui ne font pas droit à un certain pluralisme épistémologique et au « *travail normatif* »¹²² réalisé par les acteurs de terrain ne sont finalement pas aussi efficaces qu'elles prétendent l'être. L'approche positiviste du redoublement est en effet biaisée dès le départ, car elle repose sur des implicites, des postulats inadéquats. Les acteurs parties prenantes de la mise en œuvre, ceux qui sont en première ligne, les enseignants en particulier, perçoivent d'ailleurs souvent intuitivement le problème sur lequel nous mettons le doigt.

Finalement, c'est au regard du critère d'efficacité pour légitimer l'action publique que l'on peut revenir sur la question de savoir si les méta-analyses offrent une bonne ou une mauvaise abstraction. En l'occurrence, nous sommes d'avis qu'elles proposent plutôt une mauvaise abstraction dans la mesure où cette abstraction est en décalage avec les catégories de l'entendement professoral. Or, dans la mesure où ce sont eux les acteurs de la mise en œuvre de l'action publique, et qu'il faut les convaincre du bien-fondé de l'inefficacité du redoublement, on ne peut faire l'impasse sur leur point de vue. Au fond, les méta-analyses offrent au mieux un savoir incomplet. Fonder l'action publique sur ces seules données reviendrait à se bercer d'illusions et à induire les décideurs en erreur.

121. Laurent Lima, Marina Tual, « De l'étude randomisée à la classe : est-il suffisant d'avoir des données probantes sur l'efficacité d'un dispositif éducatif pour qu'il produise des effets positifs en classe ? », *Éducation et didactique*, vol. 16, n°1, 2022, p. 153-162.
122. Nicolas Dodier et Janine Barbot, « La force des dispositifs », art. cit.