Article

ROE

ECNU Review of Education
1–26

DOI: 10.1177/2096531120928086
journals.sagepub.com/home/roe
SAGE
ECNUP

Evidence-Based Education: The (Not So Simple) Case of French-Speaking Belgium

Dylan Dachet
University of Liège

Ariane Baye
University of Liège

Abstract

Purpose: Taking the case of educational reform in French-speaking Belgium (the *Pact for Educational Excellence*), this article aims to describe the process of integrating evidence-based education (EBE) into an educational reform within a country where the influence of qualitative research is important.

Design/Approach/Methods: A narrative case study was conducted to analyze and understand the stages of the ongoing educational reform process and the successive involvements of our research center in the work of this reform. Inspired by the paradigm of EBE, we initiated both the consultation processes intended to set the direction of educational decisions and programs.

Findings: According to Slavin, providing educator-friendly reviews, making available a broad range of proven education programs, and providing resources to help schools to implement proven programs are levers for implementing evidence-based reform within an educational system. This article shows that one of the three levers is fully met. For the other two, work being done on EBE in the U.S. and the UK in particular can save us time.

Corresponding author:
Ariane Baye, Centre for Analysis and Interventions in the fields of Dropout and Exclusion (AIDE), University of Liège, Liège 4000, Belgium.
Email: ariane.baye@uliege.be

Originality/Value: This article highlights concrete actions that will ensure the gradual establishment and acceptance by political and educational actors of an evidence-based reform.

Keywords

Belgium, education system reform, evidence-based

Date received: 27 November 2019; accepted: 23 April 2020

Introduction

Since the beginning of the 21st century, many political and state institutions have drawn on the results of fundamental research, particularly in the humanities and social sciences, to guide decision-making processes or support for public action (Quéré, 2017). Inspired by the paradigms of the biomedical sciences, evidence-based education (EBE) reflects this trend by providing practitioners and actors in the world of education with programs whose effectiveness has been proven in methodologically rigorous scientific studies. In French-speaking Belgium, the OECD's findings on the basis of the data from Programme for International Student Assessment (PISA) are incontestable: In addition to the fact that this part of Belgium ranks among the low-performing industrialized countries (Organisation for Economic Co-operation and Development [OECD], 2016), its education system tends to increase the differences among students according to their socioeconomic status (Demeuse, 2005; Demeuse et al., 2005; Lafontaine et al., 2019). Under pressure from these results of international surveys, there has been a sense that an in-depth reform of the education system is needed. Historically, however, French-speaking Belgium has drawn little if any inspiration from evidence-based paradigms in its reforms of the education system. The best example is probably the practice of grade retention (i.e., the practice of having students repeat a year) which, despite the vast array of evidence pointing against it since the 1970s (Allen et al., 2009; Baye et al., 2019; Galand et al., 2019; Holmes, 1990; Holmes & Matthews, 1984; Jackson, 1975; Jimerson, 2001; Lorence, 2006), has actually become more common over the last 20 years, affecting 46% of the students between grades 1 and 10 according to the PISA 2015 results (Crahay & Marcoux, 2019; Lafontaine et al., 2019).

In 2015, the double challenge of efficiency and equity led the Minister of Education to undertake a reform which, unprecedentedly for Belgium, was to be driven by the involvement of all educational stakeholders in the system (trade unions, private and public providers of schooling, parent associations, etc.), including the educational science research units of all universities and colleges within the education system.

This article aims to describe the gradual steps taken by the education system in this part of Belgium toward an EBE policy. With this in mind, we outline the history of the evidence-based

approach and of the education system in French-speaking Belgium. Next, we analyze structural reforms, putting them into perspective with the best evidence available in educational science. Finally, we describe the reform process in progress in light of the three levers for implementing an evidence-based reform in an educational system, according to Slavin (2017, 2019): (1) making a broad range of proven programs available; (2) promoting educator-friendly research reviews; and (3) providing resources to help schools to implement proven programs.

Evidence-based education

Historically, the first mention of the term "evidence-based" in the scientific literature was in 1991 (Smith & Rennie, 2014), when Gordon Guyatt outlined two opposing approaches to solving a problem situation in the biomedical field: On the one hand, reliance on the medical expertise of established personnel, which Guyatt described as the approach of the past, and on the other hand, the use of research from the scientific literature—the approach of the future (Guyatt, 1991). The following year, Guyatt et al. (1992) published their famous paper "Evidence-Based Medicine. A New Approach to Teaching the Practice of Medicine." This paper, which some people, such as Zimerman (2013), would describe as a political manifesto, called on the scientific community to make a profound change in its paradigm of clinical practice, with the aim of relying on the best evidence available from empirical studies rather than on the intuition and personal experience of professionals (Guyatt et al., 1992). This new approach, evidence-based medicine, may be defined as "the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients" (Sackett et al., 1996, p. 71). It involves combining professionals' individual clinical expertise and skills, the best evidence from the scientific literature, and patients' preferences and values (Sackett et al., 1996).

Bringing to the fore the best evidence from rigorous scientific studies in clinical decision-making thus has few precedents in the history of biomedical practice. This best evidence, although it often comes from randomized controlled trials (RCTs) and systematic reviews of the literature (Li Wan Po, 1998), may also come from studies with alternative research designs which are better suited to providing answers to the clinical question being asked by the practitioner (Sackett et al., 1996). While experimental studies are appropriate for measuring the efficacy of given treatments, cross-sectional studies may be suitable for evaluating the accuracy of a diagnostic test, and longitudinal studies on certain categories of patients can be useful for optimizing medical personnel's prognoses (Petticrew & Roberts, 2006; Sackett, 1989; Sackett et al., 1985, 1996). Furthermore, advances in the fundamental sciences such as genetics or chemistry must also be taken into account in professionals' everyday business (Sackett et al., 1996).

The use of the best evidence from empirical research by practitioners in their clinical decision-making has raised the problem of how to provide and make available to practitioners

recommendations from scientific studies. Three issues may be pointed out to explain this difficulty: (1) the quantity of studies, which has increased exponentially over the last few decades, (2) the time practitioners may devote to read it, and (3) the very purpose of these studies. (1) The scientific community has frequently observed the difficulty experienced by practitioners wishing to keep up with the scientific literature. In 1981, Sackett reckoned the literature's growth rate to be nearly 6–7% per year in the field of biomedical research. Bastian et al. (2010) estimated that 14 RCTs were being published every day in the field of biomedical research in the late 1970s. Their estimate for the number of experimental studies currently published every day is nearly 75. These figures suggest that the workload of professionals wishing to keep up with the scientific literature is unmanageable, especially when considering (2) their available time for this task. Sackett (1995) estimated, on the basis of a survey of medical specialists in several British hospitals, that practitioners have less than an hour a week to spend reading scientific articles. The last issue (3) is the final aim of the scientific papers. The survey conducted by Williamson et al. (1989) on a two-level random sample of 625 doctors points out that, according to the professionals, the scientific literature is often poorly written and poorly organized, so that transposing it to a concrete clinical application is a complex business. As Haynes (1990) noted, making connections between peer-reviewed scientific journal articles and clinical practice is increasingly difficult, because scientific articles are produced as tools for communication between researchers rather than communication with practitioners. The purpose they serve means that such articles rarely reach definite conclusions, instead presenting innovations and hypotheses in need of more extensive study. These issues led Archie Cochrane in 1979 to promote the creation of "critical summaries" of all RCTs in the medical field to support professional practice.

EBE brings this originally biomedical paradigm into the field of educational science. Although the use of experimental methods to study the human and social sciences, including education, is far from new (Cook & Campbell, 1979; Cronbach, 1957; de Landsheere, 1986), educational policies were for a very long time governed by rules, mainly state-imposed and political and financial in nature, or even by fashion, and not by scientific research (Slavin, 2016). It was not until the 1980s and the birth of the "test-based accountability" movement (Slavin, 2016) that reflections on the effectiveness of education began to influence educational policy in the U.S. This movement, which saw the introduction of rewards and sanctions for teachers and schools based on their results in standardized national tests of school performance (Hamilton et al., 2002), did not, however, always have the desired effects. Von der Embse et al. (2016) have shown, using a structural equation model to analyze data from 6,428 teachers collected on the basis of the Teacher Stress Inventory, that these policies significantly increased the general level of stress felt by teachers. This was especially so since these policies were not combined with the dissemination of effective educational programs to support teachers in achieving their targets. The literature review conducted by Herman (1999) confirmed the need to provide more evidence-based programs, since it showed that

only 20.8% of school-wide programs could be considered to be "based on strong evidence of effectiveness." This problem led to the foundation in the late 1990s of the "evidence-based practices in education" movement, brought to the fore in the U.S. by the No Child Left Behind Act (2001); this movement encouraged the use of pedagogical methods that were supported by scientific research and that would meet the increased demand for more effective teaching that had been created by accountability policies (Moran, 2004). At the same time, a similar movement emerged in the UK (Norman, 2016), in response to criticisms of the functioning and results of educational research (Davies, 1999), which rarely used scientifically rigorous (i.e., experimental) designs and produced few education programs that were interesting, usable, and reproducible for practitioners (Baye & Bluge, 2016).

This desire for reproducibility (from both a pragmatic and a financial perspective) has been simultaneously the strength and the stumbling-block of the evidence-based paradigm. Its detractors have accused it of seeking the standardization of teaching, as opposed to genuine appropriation by practitioners of empirical knowledge derived from research (Dupriez, 2015; Lessard, 2007; Saussez & Lessard, 2009; Vandenbroeck, 2018; Vandenbroeck et al., 2012). Some of these researchers (Saussez & Lessard, 2009) have preferred the term "evidence-informed education" to "evidence-based education." They have argued that, while it is true that evidence can be derived from experimental research, the contributions of other research methods in educational science should not be neglected, in order to avoid any epistemological impoverishment (Hammersley, 2001). Furthermore, as in the fields of social sciences (Petticrew, 2015; Petticrew & Roberts, 2006) and biomedical research (Grant & Booth, 2009; Lorenc et al., 2012), educational science cannot find answers to all the educational questions posed by practitioners using only RCTs or systematic reviews. For example, the effect of structural changes in educational systems can only be measured with large-scale correlational studies such as PISA. Finally, although some researchers claim that education is a more complex process (being culturally and contextually specific) than health care (Hammersley, 1997), Davies (1999) showed that evidence-based medicine, like EBE, has been equal to the challenge of studying complex realities.

The foundations had been laid of a movement which, in the years that followed, led to an increase similar to that observed in medicine (Bastian et al., 2010; Chalmers et al., 2002) in projects in the English-speaking world reviewing experimental and quasi-experimental studies of the effectiveness of educational programs. The most successful of these are *Best Evidence Encyclopedia, What Works Clearinghouse*, and the work of the *Education Endowment Foundation*.

The education system in French-speaking Belgium

In order to outline the changes that should occur to the education system in French-speaking Belgium after the major structural changes planned for the period to 2030, it is first necessary

to describe its current structure. The structure and current functioning of education in Belgium are defined by the Belgian constitutional reform of 1988, known as the Third State Reform (Parlement de la Fédération Wallonie-Bruxelles, 2011). Reflecting political and territorial changes leading to the Federation of the Belgian linguistic entities (Dutch-speaking, French-speaking, and German-speaking), it grants to the three linguistic Communities almost all competencies in terms of education (de Bouttemont, 2004). Consequently, the French-speaking Community of Belgium has a large degree of autonomy in the organization of compulsory education in its territory, although some federal laws remain applicable throughout Belgium, such as those stipulating that compulsory education should be free of charge, and that education should be compulsory from 6 to 18 years (Beckers, 2006; Demeuse, 2005).

A few current characteristics of the education system in French-speaking Belgium are worth describing, since they are at the center of the reform to come. After presenting these characteristics, we will analyze the plans concerning them in the context of the reform toward a more evidence-based approach:

- An approach of *education through skills*, as revealed by the *Socles de Compétences* (Fédération Wallonie-Bruxelles, 2013), a legal document defining the core skills which are to be mastered by the end of the first 8 years of compulsory education and on which all the educational programs must be based.
- *Pre-primary education that is not compulsory, although it is attended by 97% of 3–5-year-old children* (Fédération Wallonie-Bruxelles, 2019a), and not based on proper educational standards objectively describing what children should learn, including the skills whose acquisition should be confirmed by certification (Fédération Wallonie-Bruxelles, 2013).
- A five-cycle structure (Beckers, 2006) characterized by a *first, early choice of orientation (pre-track)* at age 12 between general and differentiated education and *a second orientation* at age 14 between general and vocational *tracks*. Academic failure and difficulties often lie behind a choice of streams considered to be for those who have failed elsewhere (Demeuse & Lafontaine, 2005; Dupriez, 2010; Mons, 2007), as the common core syllabus is common only to students with school results which are deemed satisfactory.
- *Frequent use of grade retention* in the management of student difficulties (Demeuse et al., 2005; Lafontaine et al., 2019). Average retention rates are 3% per grade in primary education and 12% per grade in secondary education (Fédération Wallonie-Bruxelles, 2018a); as a result, nearly 22.7% of students have fallen behind their peers by the end of primary education, and nearly 60% of students by the end of secondary education (Fédération Wallonie-Bruxelles, 2018a).

- A very full school calendar. Students in French-speaking Belgium spend close to *7,700 hr* (the OECD average is 6,732) *in the classroom between the ages of 7 and 14* (OECD, 2011, 2014), and holidays are not uniformly distributed: In particular, students are given a *2-month summer holiday* (Order of 27 March 2019 setting the holidays in basic and secondary education for the 2019–2020 school year).
- An *initial nonuniversity training program for elementary and lower-secondary education teachers* of 3 years for basic education and lower-secondary education (Decree of 12 December 2000 defining the initial training of teachers in basic and lower-secondary education).

Unfortunately, the education system's structure and its inherent characteristics do not seem to enable students in French-speaking Belgium to achieve satisfactory results in international surveys, with PISA 2015/2018 showing results below the average for OECD countries in science and reading and close to the average in mathematics (Lafontaine et al., 2019; Quittre et al., 2017). Worse, social determinism is particularly strong and seems to be becoming more pronounced (Monseur & Baye, 2015; Quittre et al., 2017, Lafontaine et al., 2019). Some authors go so far as to conclude: "A characteristic of the French-speaking Belgian school system seems to be emerging: students in vulnerable categories are at higher risk than in most other systems of obtaining relatively weak results" (Demeuse et al., 2005, p. 271).

The *Pact for Educational Excellence*

In 2015, at the initiative of the Minister of Education, all the educational stakeholders of French-speaking Belgium, from researchers to policymakers, and including the providers of schooling, trade unions, parents' representatives, and others, came together primarily with a view to reforming the education system in French-speaking Belgium so that it could offer an effective response in the areas of performance and equity where it had previously failed. The various stakeholders were invited to form thematic working groups responsible for synthesizing all the available data on the current state of the education system in relation to certain predefined themes. The findings of these different working groups were condensed by the *Groupe central* to produce a comprehensive analysis of the current state of the education system (Fédération Wallonie-Bruxelles, 2015). It was clear from this analysis that a structural educative reform was needed. This was the *Pact for Educational Excellence*.

A scientific steering committee was created for the reform, on which universities were represented. The representatives of the University of Liège argued for more programs and recognition of the findings of rigorous research, that is, the introduction of evidence in the context of the reform. Our research center was then invited to produce a report introducing the unknown concept of EBE

to the scientific and political unit in charge of the ongoing reform (the *Groupe central*). Based on Slavin's work (2002, 2007), we defined EBE as a process of change based on rigorous scientific research in order to influence education policies and practices. We also anticipated various fears or misunderstandings concerning EBE (Baye & Bluge, 2016): the confusion between EBE and over-standardization, the reduction of educational outcomes to a restricted range of measurable outcomes, and the fear of an epistemological impoverishment. We argued that, in our education system, EBE was certainly the research paradigm that had been the less encouraged in the past.

Unexpectedly, the Baye & Bluge (2016) report had a considerable impact, with some of its findings being included in *L'avis n°3 du Groupe central* (the recommendations of the administrative/political unit preparing the reform) (Fédération Wallonie-Bruxelles, 2017). *L'avis n°3 du Groupe central* defines EBE as a way to provide teachers and schools with tools and programs validated by research evaluating interventions' effects. It acknowledges that EBE is virtually nonexistent in our education system and proposes that more account should be taken of research results, as defined from an EBE perspective. *L'avis n°3 du Groupe central* acknowledges the dynamic essence of EBE: taking new evidence into account to improve practices. Moreover, it introduces the idea of piloting validated programs from aboard in order to scale them up once evidence of their efficacy has been collected, with incentives for schools using such programs. From this point onward, EBE has colored the *Pact for Educational Excellence*.

As we hope that this explicit reference to EBE has impacted the structural changes planned in the reform, we will start our analysis by putting the structural reforms into perspective with the best evidence available in educational science. We will then consider how the process of gathering, selecting, and validating educational tools and programs within the *Pact* contributes to the evidence-based nature of the reform according to Slavin's levers (2017, 2019).

Are structural changes supported by educational research?

The *Pact* thus set out plans for a set of structural changes to the education system. Table 1 summarizes the most important structural aspects of the reform. In the third column of the table, we compare the planned structural changes with the *best available* evidence from research in education that indicates whether their implementation is justified. The purpose of doing this is to show how the *Pact* and its specific structural proposals relate to the concept of evidence-based policy. As structural changes cannot always be assessed via experimental design, we used mainly correlational studies, which are better suited to providing this kind of evidence (Sackett et al., 1996). Details and conclusions from the studies referenced in Table 1 are provided in Online Appendix 1.

The comparison of seven planned structural changes with the best scientific evidence (Column 3) makes it clear that some of the proposals are consistent with scientific evidence of various kinds. Evidence from correlational studies indicates that expanding the compulsory common core from

Table 1. Comparison of the situation before and the planned situation after the *Pact for Educational Excellence* and what is known about the educational effectiveness of the proposed structural changes.

Situation in French-speaking Belgium		
Before the *Pact*	Intended after the *Pact*	Best available evidence
A skills-based approach defined by the *Socles des Compétences* (Fédération Wallonie-Bruxelles, 2013)	An approach based on skills and definition of knowledge (see the *Charter of Educational Standards* approved by the government in 2017)	The skills-based approach is an example of a large-scale reform lacking in scientific research as shown in recent literature reviews (Lassnigg, 2015; Morcke et al., 2013).
Non-compulsory pre-primary education without curriculum	Compulsory pre-primary education from age 5 and creation of suitable educational standards (Fédération Wallonie-Bruxelles, 2017)	International surveys (PISA) show that a pre-primary education affects achievement at the age of 15 even when socioeconomic status is accounted for (OECD, 2010).
Common core until 12–14 years (Demeuse & Lafontaine, 2005)	Common core, expanded and extended from 3 to 15 years (Fédération Wallonie-Bruxelles, 2017; Pacte pour un Enseignement d'Excellence, 2017)	International surveys (PISA) indicate that the most equitable education systems are also those which delay the initial orientation of students the longest (Demeuse & Baye, 2008; Monseur & Lafontaine, 2012).
Early academic orientation of students according to their achievement, their failures, and the difficulties they have experienced in general education (Canzittu, 2018)	Orientation of students based on the orientation approach model (Canzittu, 2018; Fédération Wallonie-Bruxelles, 2017)	Despite extensive theoretical work (Canzittu & Demeuse, 2017; Franquet, 2010), there have been few attempts to validate this approach through rigorous research.
Frequent use of grade repetition (Fédération Wallonie-Bruxelles, 2018a)	Objective of reducing the grade repetition rate by 50% by 2030 (Fédération Wallonie-Bruxelles, 2019b)	None of the exhaustive and critical reviews of the literature on grade repetition have succeeded in identifying any positive mean effects (Allen et al., 2009; Baye et al., 2019; Holmes, 1990; Holmes & Matthews, 1984; Jackson, 1975; Jimerson, 2001; Lorence, 2006).

(continued)

Table 1. (continued)

Situation in French-speaking Belgium		
Before the *Pact*	Intended after the *Pact*	Best available evidence
Unbalanced school calendar (Arrêté Vacances et Congés dans l'Enseignement, 2019)	School calendar alternating 7 weeks of school and 2 weeks of holiday; summer holidays reduced and other holiday periods extended (Fédération Wallonie-Bruxelles, 2017; Fédération Wallonie-Bruxelles, 2018b)	The new school calendar, which is the subject of extensive theoretical work (Testu & Fontaine, 2001; NECTL, 2005) and work based on children's biological and psychological needs (FAPEO, 2008), seems likely to be welcomed by the population (Mouraux, 1992). Although a study of its economic impact has been carried out (Matthys & Penxten, 2014), evidence on new calendar programs, such as year-round school calendars, is insufficient (Finnie et al., 2019).
Three years of initial training (180 European Credit Transfer and Accumulation System (ECTS)) for elementary and lower-secondary teachers in "normal schools" (Decree of 12 December 2000 defining the initial training of teachers in basic and lower-secondary education)	Four years of initial training for elementary and lower-secondary teachers with a university component (180 ECTS + 60 univ. ECTS) (Decree of 5 March 2019 defining the initial training of teachers)	Studies suggest that enhancing teachers' skills leads to an improvement in students' academic performance (Rivkin et al., 2005; Rockoff, 2004). Such findings support the improvement of initial teacher training.

the age of 5 to 15 is consistent with the aim of reducing social inequalities (Demeuse & Baye, 2008; Monsieur & Lafontaine, 2012). Correlational studies also indicate that improving teachers' skills improves students' performance (Rivkin et al., 2005; Rockoff, 2004) . . . assuming reasonably that extending initial teacher training by one year will produce more skilled practitioners.

The other four planned structural changes deserve more comment, either because the amount of rigorous research available on a topic could have justified a much more ambitious reform—such as the grade retention policy—or because, conversely, the lack of rigorous research should lead to much more caution before a given change is implemented widely.

This is the case with the skills-based approach. It remains central to the definition of all standards for compulsory education in French-speaking Belgium. However, specific systematic reviews such as that by Morcke et al. (2013), based on eight studies of the performance of medical students, or the literature review by Lassnigg (2015), based on 33 studies of the effectiveness of competence-based education, have shown the lack of rigorous research into its educational effectiveness. Yet this fashionable approach seems to have maintained its appeal, although the swinging back of the pendulum is also observable with the reintroduction of knowledge in educational standards.

In the same way, the "orientation approach model" (Canzittu, 2018; Canzittu & Demeuse, 2017) lies behind the intention of restoring the image of the academic and vocational streams in schools. Although it is based on educational intentions and principles that may be effective (Dupont et al., 2002; Franquet, 2010), this model lacks the support of a rigorous scientific study with an experimental research design.

A last example is the one of the school calendar change. While its intentions are laudable and it can be said to have psychobiological foundations (FAPEO, 2008; Mouraux, 1992; Testu & Fontaine, 2001), the reform of the school calendar is not based on any evidence of its educational effectiveness in terms of academic achievement (Finnie et al., 2019).

By contrast, given the widespread use of grade retention in French-speaking Belgium, the objective of reducing it by 50% is regrettably inadequate. Although there is some awareness of our education system's anomalous position as a world leader in terms of grade repetition, it will still be possible for teachers to use this practice in the common core grades (Fédération Wallonie-Bruxelles, 2017). The *Groupe central* (Fédération Wallonie-Bruxelles, 2017, p. 5) states in this regard that "although it is out of the question to ban grade repetition throughout the common core, it must become the exception and not the rule in view of its ineffectiveness." The system clearly favors the second option presented by Perrenoud (1996), who describes the tension between an authoritarian, top-down policy strategy (in which direct and comprehensive policymaking is initially greeted with suspicion and outrage in the court of public opinion, but eventually leads to a change in popular attitudes) and a soft, bottom-up political strategy (involving the dissemination of research results, the modification of working relations between teachers, an incentivization policy and other measures) with respect to lowering grade repetition rates.

How does the process of making programs available contribute to the evidence-based nature of the reform?

The *Pact* set up eight working groups made up of didactics specialists and educational scientists from the "normal schools" and universities. These eight working groups were responsible for providing teachers with tools/programs recommended by these recognized experts in our education

system in different fields (1—Modern languages; 2—French and Latin; 3—Artistic sensibility and expression; 4—Mathematics, science, and physical geography; 5—Technical subjects, technology, computer science, and manual work; 6—Humanities, social sciences, philosophy, and citizenship; 7—Physical education, well-being, and health; 8—Digital education). The aim is to keep teachers informed about existing tools/programs, in order to help them as the common core is introduced. The substantial work involved in producing an inventory of tools and programs began in 2016. As a result of our report on EBE (Baye & Bluge, 2016), we were contacted to provide support to the working groups set up the previous year. Our research center was asked to perform a descriptive analysis of the research designs behind all the tools identified by the eight working groups in order to study their scientific validity, and ultimately their effectiveness.

At the outset of this work, we were expecting to come up with various precise criteria for identifying tools/programs based on studies with experimental or quasi-experimental designs, so that only the most rigorous could be kept. We were amazed to find that out of the 425 tools/programs listed by the working groups in 2017, only 61 had been previously tested in studies conducted by educational researchers. Moreover, of these 61 tools/programs, just 13, or 3% of the total, had been tested according to an experimental design.

The question now arose of how to respond to this state of affairs. Traditionally, evidence-based models primarily advocate empirical research of the experimental type. However, as we saw in the historical review of evidence-based medicine, this paradigm relies above all on using the best evidence currently available on a given issue (Sackett et al., 1996). The low rate of experimental studies conducted in contexts comparable to that of French-speaking Belgium and with pedagogical and didactic tools/programs consistent with its educational programs ruled out a hierarchical (Bouffard, 2012; Laurent et al., 2009) or even a pyramidal (Murad et al., 2016; Pageau, 2016) approach to evidence. Such an approach would have led us to exclude almost 97% of the tools/programs proposed for French-speaking educational professionals in Belgium! In any case, a hierarchical approach is a commonly criticized feature of EBE, and that criticism had also been voiced by some of the researchers in the working groups. Moreover, although such a high exclusion rate would not have been problematic scientifically, politically, it was a real issue. It was therefore essential to accept methodological diversity, and standards that were different from those defined by organizations such as *Education Endowment Foundation*, *What Works Clearinghouse*, or *Best Evidence Encyclopedia* and that would be better adapted to the relatively recent nature of experimental research paradigms and evidence-based models in French-speaking Belgium. We therefore created a typology (Table 2) that respected methodological heterogeneity, while hoping to convince certain didactics specialists of the value of experimental research designs in a context of positive collaboration. This decision made sense, as the didactics specialists were our best potential allies, given that they are responsible for a large part of teacher training and thus have considerable influence.

Table 2. Inclusive typology of research designs adapted to the work of the *Pact*.

Type of validation	Description of the study designs
Validation by *experts* with specific roles (practitioners, researchers, designers, etc.) and *corroboration* (by empirical data or theoretical principles)	This type of validation is usually based on an analysis of the tool, its principles, and its relevance (without any collection and systematic analysis of data). We differentiate between: (1) *Individual expertise*: This type of validation comes from the ability of an expert to work "with the tools of their trade and immersed in the ideas of their intellectual communities" (Thomas, 2004, p. 3). (2) *Panel of experts*: This type of validation comes from "independent specialists, recognised in at least one of the fields addressed by the program under evaluation. The panel specialists arrive at conclusions and recommendations through consensus" (Evaluation Unit DEVCO, 2015). It is important to back up the experts' validation with *bibliographical references* to data from empirical research or psychopedagogical or didactic principles which form a basis for the proposal of a tool/program. This information makes it possible to claim that, even if the proposed tool/program has not been the subject of data collection among students in French-speaking Belgium, its recommendation by one or more experts is consistent with the current state of knowledge. Such references are elements in the *corroboration* (Thomas, 2004) of the tool/program.
Validation by one or more *case studies* using qualitative and/or quantitative data	This type of validation looks at all systematic studies of the tool, its characteristics, use, and impact. It includes: (1) *Case studies* using qualitative methods to describe one or more cases considered to be of interest (Delvenne, 1999). These "make use of various methods such as observation, semi-directive interviews and one or other of the content analysis techniques" (Hamel, 1997, p. 103) and contribute

(continued)

Table 2. (continued)

Type of validation	Description of the study designs
	"to research of the exploratory type and the understanding of factors that are difficult to measure" (Alexandre, 2013, p. 27). (2) *Cross-sectional studies* using quantitative methods to describe a population on a one-off basis (Debout, 2012) in terms of the frequency of a variable and the characteristics associated with that variable (Delvenne, 1999; Mercklé, 2010).
Validation by one or more *correlational studies* or a *longitudinal study*	This type of validation refers to two types of statistical study: (1) *Correlational studies*, which seek to model the relations between some variables, often studied in their natural environment (Debout, 2012) and on the basis of multi-level schemes. (2) *Longitudinal studies*, which collect quantitative data based on repeated measurements (Shaughnessy & Zechmeister, 1990; Sommer & Sommer, 1991) and measure the effect over time (Newby, 2010) of different treatments and/or conditions (Delvenne, 1999), without using a control group (Petticrew & Roberts, 2006). Pretest/posttest studies without a control group are included in this category.
Validation by one or more *experimental* and *quasi-experimental studies* (including a control group and systematic testing of the tool/program)	This type of validation is based on the use of research designs that allow a comparison to be made between one or more experimental treatments and one or more control treatments. A distinction is made between: (1) *RCTs*: These use a comparative experimental research design to measure the impact of one (or more) tool/program by comparing the results of an experimental group—in which students are taught using the tool/program—with those of a control group in which students are taught in the usual way. In an RCT, the distribution of subjects (students, teachers, or schools) into the experimental or

(continued)

Table 2. (continued)

Type of validation	Description of the study designs
	control group is strictly random (Slavin, 2007). The random assignment of subjects—especially if their number is large—ensures the equivalence of the groups before the tool/program is implemented and avoids selection bias (Slavin, 2007). (2) *Quasi-experimental studies* (with matching): These have a comparative design and measure the impact of one (or more) tool/program by comparing the results of an experimental group—in which students are taught using the experimental tool/program—with those of a control group taught in the usual way. In these studies, the control group must be defined so as to be fully comparable with the experimental group on as many dimensions as possible, and especially with regard to the behaviors or aptitudes under evaluation (Slavin, 2007). The timing of the control group's selection is important: Studies have shown that when control groups are selected after the start of the implementation of the tools/programs (post hoc selection), the results often favor the experimental group (Cheung & Slavin, 2016).

Note. RCT = randomized controlled trial.

This typology has two distinct purposes. The first goal was to make it possible to produce an inventory of the tools/programs selected by the working groups that would take account of their considerable methodological diversity and to classify each program according to its level of evidence. Second, this typology allowed us to define the minimum standards for each of the methodological approaches chosen, including qualitative approaches. It also provided a framework for the research work carried out by the working groups which, as well as compiling an inventory of current tools/programs, also had the task of studying their implementation in the classroom. Some of the proposed tools/programs had never been the subject of a classroom implementation study. For example, our typology defines a number of methodological quality standards for qualitative case studies reporting on a tool's implementation in different classroom contexts. Through this typology, we asked that any tested tools/programs should be implemented in at least

two classrooms from environments with differing characteristics (school's socioeconomic status, teaching network, extent of teachers' experience, etc.).

Three central ideas guided the creation of this typology. First, different types of research complement each other by answering questions of a different nature (Petticrew & Roberts, 2006). Second, the output of the working groups was valuable in that, at the very least, it identified 425 educational and didactic tools/programs evaluated by a panel of experts consisting of researchers, didactics specialists, and professors from the various universities and university colleges of French-speaking Belgium; it thus enabled practitioners to select tools/programs for use in the classroom on the basis of a preliminary form of validation in the selection of tools/programs instead of relying entirely on their intuition (Thomas, 2004). Third, nonexperimental studies can at least reveal that it is worth carrying out more in-depth and rigorous studies on the actual implementation and effect of tools/programs in a real classroom setting, along similar lines to the "promising programs/projects" in English-speaking countries (Education Endowment Foundation, 2019).

Our first argument in favor of this typology, concerning the complementarity between different types of research by virtue of their differing focuses, was inspired by the work of Petticrew and Roberts (2006), who created a model linking different research methods with the prototypical research questions they seek to answer. As this work was originally intended for the social sciences, we adapted it to the more specific field of educational science and to the context of the work of the *Pact*. This adaptation is presented in Table 3. The different column headings present all the research designs in our typology. Each row considers a set of questions that may be asked by researchers wishing to study the implementation of an educational program and/or tool in a classroom context. As well as questions about the effectiveness of the program, researchers often seek to understand how it is actually implemented, find out what the teachers who use it think of it, study the links between traditional classroom practices, and so on. In the body of the table, in the same way, as in the work of Petticrew and Roberts (2006), a varying number of plus signs indicate the degree of match between the research design and the questions formulated.

According to Slavin, "for evidence-based reform to prevail, three conditions must exist." The first is to make available "a broad range of proven programs in key areas of education, at every grade level" (Slavin, 2019, p. 2). In the *Pact*, an inventory of existing programs in French-speaking Belgium has been completed. It emerges from this inventory that we are far from meeting the evidence-based criteria, since almost no experimental study is available in our context. However, the inclusive nature of the typology and close collaboration with the didactics specialists have drawn attention to the experimental methodology and shown its added value. Therefore, for the first lever, we would conclude that we are at the very beginning of the path.

Table 3. Complementarity among methods, focuses, and research questions.

Type of validation		Panel of experts	Case study	Cross-sectional study	Correlational study	Longitudinal study	(Quasi-)experimental study
Research question	Aim	Selective	Descriptive		Relational		Comparative
Which programs have been studied in depth?		+	++		++	++	+++
Which programs are promising?		+	++		++	++	+++
How well do professionals receive a program?			+++	+++			
How is the program implemented by professionals?			++			++	+++
What adaptations of the program are made by professionals in the classroom?			++			++	+++
How close is the match among target group, objective and program?			+	+	++	+	+++
What measurement tools are used?			+	+++			
What is the link between classroom practices and academic performance?					++	+	+++
How effective is a structural change in an education system?					++	+	+++
How has a category of the school population changed over time?						++	+++
How effective is an educational program?					+		+++
What is the cost-benefit ratio?					+		+++

Given this conclusion on the first lever, it is obvious that we do not have enough experimental studies to carry out educator-friendly reviews, the second criterion of Slavin (2017, 2019). We wanted to initiate a change in this direction by selecting two priority areas for improvement: reading and dropout. For each of these key areas, we have selected a program from abroad. These programs are being assessed with a quasi-experimental design. Furthermore, we have selected systematic reviews from the *Best Evidence Encyclopedia* in order to translate and make available educator-friendly reviews of research from aboard. Incidentally, we wish to stress that the infrequent use of experimental research in the selection of teaching tools/programs by the working groups is not solely due to the relatively recent nature of experimental research paradigms in French-speaking Belgium: The problem is also linked to the difficulties encountered by French-speaking educational actors in staying abreast of the English-language scientific literature, full though that literature is of studies that use the evidence-based paradigm. As well as coping with the inherent complexity of scientific articles (Haynes, 1990; Williamson et al., 1989), these actors have to overcome the barrier of reading in English, which is not one of Belgium's three national languages, but whose primacy among scientific articles with a high impact factor is beyond dispute (Hamel, 2013). In support of this observation, we counted the number of scientific references to the English-language literature in the first year of work: They represented less than 15% of the total!

We are more enthusiastic about the use of Slavin's third lever in our reform, that is, providing resources to help schools to implement proven programs. *L'avis n°3 du Groupe central* explicitly mentions such resources:

> In line with the report by Baye and Bluge et al. 2016, widespread promotion of experimentation in schools, associated with a specific and recurrent budget for a selection of projects submitted to rigorous evaluation. This budget is justified by (1) the poor existing resources for policy evaluation and experimentation in French-speaking Belgium, (2) the high potential impact in the long term of experimentation and a systematic evaluation of public policies, (3) the need to assess a large number of initiatives planned in the *Pacte* to ensure their effectiveness before they are widely introduced. This corresponds to an estimated budget of 2 million euros. (Fédération Wallonie-Bruxelles, 2017, p. 44)

Conclusion

One of the many ambitions of the reform of the *Pact for Educational Excellence*, work on which began in 2015, is to progressively integrate evidence-based recommendations and programs into the education system of the French-speaking Community of Belgium, which is currently struggling to meet the achievement and equity goals of schools in the 21st century.

While we welcome such a change in the education system as experimental researchers, it remains the case that this *Pact* with its evidence-based tendencies is not yet up to standard according to Slavin's criteria (2017, 2019). First, experimental research on effective programs

is still too rare in French-speaking Belgium. On the positive side, we can draw inspiration from (or even importing) programs from the English-speaking world which have been tried and tested at various regional or national levels. However it is not enough to guarantee their validity in the French-speaking context. We should note that since the report on EBE (Baye & Bluge, 2016), a first round of public funding has been allocated to experimental research to validate educational programs. Programs such as Parler Apprendre Réfléchir Lire Ensemble pour Réussir (Zorman et al., 2015) and School-Wide Positive Behavior Supports (Horner & Sugaï, 2002; Horner & Sugai, 2015) are currently being studied experimentally by researchers from the University of Liège.

Moreover, we can benefit from recent scientific literature in English. Of course, extensive translation work is necessary to make articles published in the field of educational science accessible to educational actors. This work is under progress.

At this point, we may conclude that the *Pact*, due in particular to the many different actors involved in it, tends to adopt a compromise position between the pure evidence-based paradigm and the considerations of professional development traditionally advocated by the didactics specialists. By its definition, its structure, and its proposals, the *Pact* pays particular attention to reforming both the curricula and the structures of the education system. We also wish to underline the desire to bridge the gap between practitioners and researchers by funding research conducted in schools in collaboration with teachers. Finally, the collaboration of all educational stakeholders in French-speaking Belgium, including parents, is an innovative and valuable feature of the work of the *Pact*. It has made it possible: (1) to initiate an interactional process among researchers, the education authorities, and practitioners; (2) to identify promising educational programs in French-speaking Belgium which correspond to both the recommendations of researchers and the standards and curricula; (3) to set up panels of subject experts which can be used both in the evaluation of educational programs and in their validation; and (4) to have didactics specialists and researchers in educational science jointly define minimum methodological standards for all categories of research. In the *Pact* process, we see two major lines of progress toward EBE in French-speaking Belgium. First, it has provided a unique opportunity for didactics specialists and quantitative researchers in education to work together, to learn from each other, and to make steps together toward the validation of replicable programs. Second, this is the first time that an official text has recognized the need to pilot, validate, and scale up programs according to their efficacy. This legal basis makes the French-speaking Belgium one of the few education systems that officially recognizes EBE. We expect that this legal basis will support research and practice based on evidence in education.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Supplemental material

Supplemental material for this article is available online.

References

Alexandre, M. (2013). La rigueur scientifique du dispositif méthodologique d'une étude de cas multiple. *Recherches Qualitatives*, *32*(1), 26–56. http://www.recherche-qualitative.qc.ca/documents/files/revue/edition_reguliere/numero32(1)/rq-32-1-Alexandre.pdf

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*(4), 480–499. https://doi.org/10.3102/0162373709352239

Arrêté Vacances et Congés dans l'Enseignement: arrêté du Gouvernement de la Communauté française fixant les vacances et congés dans l'enseignement fondamental et secondaire pour l'année scolaire 2019–2020. (2019, March 27). *Moniteur belge*, 46340.

Bastian, H., Glasziou, P., & Chalmers, L. (2010). Seventy-five trials and eleven systematic reviews a day: How will we ever keep up? *PLoS Medicine*, *7*(9), 1–6. https://doi.org/10.1371/journal.pmed.1000326

Baye, A., & Bluge, V. (2016). *L'éducation fondée sur des preuves*. Pacte pour un Enseignement d'Excellence.

Baye, A., Dachet, D., & Crahay, M. (2019). Faut-il faire redoubler? In M. Crahay (Ed.), *Peut-on lutter contre l'échec scolaire?* (4th ed., pp. 155–230). de Boeck supérieur.

Beckers, J. (2006). *Enseignants en communauté française de Belgique: Mieux comprendre le système, ses institutions et ses politiques éducatifs pour mieux situer son action*. de Boeck.

Bouffard, M. (2012). Pour une pratique fondée sur des données probantes "Evidence Based Practice" (EBP) en activité physique adaptée: Perspectives et enjeux. *La nouvelle revue de l'adaptation et de la scolarisation*, *58*, 77–88. https://www.cairn.info/revue-la-nouvelle-revue-de-l-adaptation-et-de-la-scolarisation-2012-2-page-77.htm

Canzittu, D. (2018). *Orientation scolaire et approche orientante : les politiques et dispositifs d'orientation dans les systèmes scolaires*. Conseil National d'Évaluation du Système Scolaire (CNESCO).

Canzittu, D., & Demeuse, M. (2017). *Comment rendre une école réellement orientante?* de Boeck.

Chalmers, I., Hedges, L. V., & Cooper, H. (2002). A brief history of research synthesis. *Evaluation & the Health Professions*, *25*(1), 12–37. https://doi.org/10.1177/0163278702025001003

Cheung, A. C. K., & Slavin, R. E. (2016). How methodological features affect effect sizes in education. *Educational Researcher*, *45*(5), 283–292. https://doi.org/10.3102/0013189x16656615

Cochrane, A. L. (1979). 1931–1971: A critical review, with particular reference to the medical profession. In G. Teeling Smith & N. Wells (Eds.), *Medicines for the year 2000* (pp. 1–11). Office of Health Economics.

Cook, T. D., & Campbell, D. T. (1979). *Quasi experimentation*. Rand McNally.

Crahay, M., & Marcoux, G. (2019). Mais pourquoi les taux d'échec varient-ils autant selon les systèmes éducatifs? Est-il possible de les faire diminuer là où ils sont élevés? In M. Crahay (Ed.), *Peut-on lutter contre l'échec scolaire?* (4th ed., pp. 53–110). de Boeck.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*(11), 671–684. https://doi.org/10.1037/h0043943

Davies, P. (1999). What is evidence-based education? *British Journal of Educational Studies*, *47*(2), 108–121. https://doi.org/10.1111/1467-8527.00106

de Bouttemont, C. (2004). Le système éducatif belge. *Revue Internationale d'Education de Sèvres*, *37*, 1–10. https://doi.org/10.4000/ries.1466

de Landsheere, G. (1986). *La recherche en éducation dans le monde*. Presses Universitaires de France.

Debout, C. (2012). Méthodologie quantitative et études descriptives simples. *Soins*, *768*, 55–60. https://doi.org/10.1016/j.soin.2012.07.017

Decree defining the initial training of teachers: décret de la Communauté française définissant la formation initiale des enseignants. (2019, March 5). *Moniteur belge*, 46261.

Decree defining the initial training of teachers in basic and lower secondary education: décret de la Communauté française définissant la formation initiale des instituteurs et des régents. (2000, December 12). *Moniteur belge*, 25501.

Delvenne, C. (1999, April 5). *Définition des principaux types d'études*. Bibliothèque de la Faculté de Médecine de l'Université de Liège. http://www.ebm.lib.ulg.ac.be/prostate/typ_etud.htm

Demeuse, M. (2005). La marche vers l'équité en Belgique francophone. In M. Demeuse, A. Baye, M.-H. Straeten, J. Nicaise, & A. Matoul (Eds.), *Vers une école juste et efficace* (pp. 191–216). de Boeck.

Demeuse, M., & Baye, A. (2008). Indicateurs d'équité éducative. Une analyse de la ségrégation académique et sociale dans les pays européens. *Revue Française de Pédagogie*, *165*, 91–103. https://doi.org/10.4000/rfp.1140

Demeuse, M., & Lafontaine, D. (2005). L'orientation scolaire en communauté française de Belgique. *Revue Internationale d'Éducation de Sèvres*, *38*, 35–52. https://doi.org/10.4000/ries.1467

Demeuse, M., Lafontaine, D., & Straeten, M.-H. (2005). Parcours scolaires et inégalités de résultats. In M. Demeuse, A. Baye, M.-H. Straeten, J. Nicaise, & A. Matoul (Eds.), *Vers une école juste et efficace* (pp. 259–273). de Boeck.

Dupont, P., Gringras, M., & Marceau, D. (2002). *Implantation d'une approche orientante. Des conditions idéales*. Université de Sherbrooke.

Dupriez, V. (2010). *Séparer pour réussir? Les modalités de groupement des élèves*. Organisation des Nations Unies pour l'éducation la science et la culture (UNESCO).

Dupriez, V. (2015). *Peut-on réformer l'école? Approches organisationnelle et institutionnelle du changement pédagogique*. de Boeck.

Education Endowment Foundation. (2019). *Promising projects: EEF-funded projects which have shown promising results when trialled*. https://educationendowmentfoundation.org.uk/tools/promising/#closeSignup

Evaluation Unit DEVCO. (2015, January 19). *Panel d'experts*. https://europa.eu/capacity4dev/evaluation_guidelines/wiki/panel-dexperts

Fédération des Associations de Parents de l'Enseignement officiel. (2008). *Le sens du rythme. Rythmes scolaires, biologiques et psychologiques de l'enfant et de l'adolescent*. Les analyses de la FAPEO.

Fédération Wallonie-Bruxelles. (2013). *Les Socles de compétences*. Enseignement.be.

Fédération Wallonie-Bruxelles. (2015). *Contribuer au diagnostic du système scolaire en FWB: Rapport à la vice-présidente, ministre de l'education, de la culture et de l'enfance*. Pacte pour un Enseignement d'Excellence.

Fédération Wallonie-Bruxelles. (2017). *Avis n°3 du Groupe central*. Pacte pour un Enseignement d'Excellence.

Fédération Wallonie-Bruxelles. (2018a). *Les indicateurs de l'enseignement*. Enseignement.be

Fédération Wallonie-Bruxelles. (2018b). *Les nouveaux rythmes scolaires annuels dans l'intérêt des élèves?* Pacte pour un Enseignement d'Excellence. http://www.pactedexcellence.be/wp-content/uploads/2018/07/Rapport_Table10_02.pdf

Fédération Wallonie-Bruxelles. (2019a). *Les indicateurs de l'enseignement*. Enseignement.be.

Fédération Wallonie-Bruxelles. (2019b, April 12). *Accord du gouvernement sur le code qui balise le nouveau tronc commun*. Pacte d'Excellence. http://www.pactedexcellence.be/index.php/2019/04/12/accord-du-gouvernement-sur-le-code-qui-balise-le-nouveau-tronc-commun/

Finnie, R. K. C., Peng, Y., Hahn, R. A., Johnson, R. L., Fielding, J. E., Truman, B. I., Muntaner, C., Fullilove, M. T., & Zhang, X. (2019). Examining the effectiveness of year-round school calendars on improving educational attainment outcomes within the context of advancement of health equity: A community guide systematic review. *Journal of Public Health Management and Practice*, *25*(6), 590–594. https://doi.org/10.1097/PHH.0000000000000860

Franquet, A. (2010). *L'approche Orientante. En route vers la réussite scolaire et professionnelle des élèves. Fascicule de présentation*. Université de Mons.

Galand, B., Lafontaine, D., Baye, A., Dachet, D., & Monseur, C. (2019). Le redoublement est inefficace, socialement injuste et favorise le décrochage scolaire. *Les Cahiers des Sciences de l'Education*, *38*, 1–29. http://www.aspe.ulg.ac.be/Files/cahiers_aspe_redoublement.pdf

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*, *26*(2), 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Guyatt, G. H. (1991). Evidence-based medicine. *American College of Physicians Journal Club*, *114*(2), A16. https://doi.org/10.7326/ACPJC-1991-114-2-A16

Guyatt, G., Cairns, J., Churchill, D., Cook, D., Haynes, B., Hirsh, J., Irvine, J., Levine, M., Levine, M., Nishikawa, J., Sackett, D., Brill-Edwards, P., Gerstein, H., Gibson, J., Jaeschke, R., Kerigan, A., Neville, A., Panju, A., Detsky, A., … Tugwell, P. (1992). Evidence-based medicine: A new approach to teaching the practice of medicine. *Journal of the American Medical Association*, *268*(17), 2420–2425. https://doi.org/10.1001/jama.1992.03490170092032

Hamel, J. (1997). *Etude de cas et sciences sociales*. L'Harmattan.

Hamel, R. E. (2013). L'anglais, langue unique pour les sciences? Le rôle des modèles plurilingues dans la recherche, la communication scientifique et l'enseignement supérieur. *Synergie Europe*, *8*, 53–66. https://gerflint.fr/Base/Europe8/Hamel.pdf

Hamilton, L. S., Stecher, B. M., & Klein, S. P. (2002). *Making sense of test-based accountability in education*. Rand Corporation.

Hammersley, M. (2001, September). *Some questions about evidence-based practice in education*. Paper presented at the symposium on Evidence-Based Practice in Education at the Annual Conference of the British Educational Research Association, Leeds, England.

Hammersley, M. (1997). Educational research and teaching: a response to David Hargreaves'TTA lecture. *British Educational Research Journal*, *23*(2), 141–161. https://doi.org/10.1080/0141192970230203

Haynes, B. (1990). Loose connections between peer-reviewed clinical journals and clinical practice. *Annals of Internal Medicine*, *113*(9), 724–728. https://doi.org/10.7326/0003-4819-113-9-724

Herman, R. (1999). *An educator's guide to schoolwide reform*. Educational Research Service.

Holmes, C. T. (1990). Grade level retention effects: A meta-analysis of research studies. In L. A. Shepard & M. L. Smith (Eds.), *Flunking grades. Research and policies on retention* (pp. 16–33). Falmer Press.

Holmes, C. T., & Matthews, K. M. (1984). The effects of nonpromotion on elementary and junior high school pupils: A meta-analysis. *Review of Educational Research*, *54*(2), 225–236. https://doi.org/10.3102/00346543054002225

Horner, R. H., & Sugai, G. (2002). The evolution of discipline practices: School-wide positive behaviour supports. *Child & Family Behavior Therapy*, *24*(1–2), 23–50. https://doi.org/10.1300/J019v24n01_03

Horner, R. H., & Sugai, G. (2015). School-wide PBIS: An example of applied behavior analysis implemented at a scale of social importance. *Behavior Analysis in Practice*, *8*(1), 80–85. https://doi.org/10.1007/s40617-015-0045-4

Jackson, G. B. (1975). The research evidence on the effects of grade retention. *Review of Educational Research*, *45*(4), 613–635. https://doi.org/10.2307/1170067

Jimerson, S. R. (2001). Meta-analysis of grade retention research: Implications for practice in the 21st century. *School Psychology Review*, *30*(3), 420–437. http://web.a.ebscohost.com/ehost/detail/detail?vid=0&sid=0a25bece-734a-4c84-87ea-e42d8069bf5b%40sessionmgr4007&bdata=JnNpdGU9ZWhvc3QtbGl2ZQ%3d%3d#AN=5508140&db=aph

Lafontaine, D., Baye, A., & Monseur, C. (2019). Redoublement, efficacité et équité: l'éclairage des enquêtes internationales. In M. Crahay (Ed.), *Peut-on lutter contre l'échec scolaire?* (4th ed., pp. 110–154). de Boeck.

Lafontaine, D., Bricteux, S., Hindryckx, G., Matoul, A., & Quittre, V. (2019). *Performances des jeunes de 15 ans en lecture, mathématiques et sciences: Premiers résultats de PISA 2019 en Fédération Wallonie-Bruxelles*. Université de Liège.

Lassnigg, L. (2015). Competence-based education and educational effectiveness: A critical review of the research literature on outcome-oriented policy making in education. *Sociological Series*, *111*, 1–41. https://doi.org/10.1007/978-3-319-41713-4_31

Laurent, C., Baudry, J., Berriet-Solliec, M., Kirsch, M., Perraud, D., Tinel, B., Trouvé, A., Allsopp, N., Bonnafous, P., Burel, F., Carneiro, M. J., Giraud, C., Labarthe, P., Matose, P. M., & Ricroch, A. (2009). Pourquoi s'intéresser à la notion d' "Evidence-based policy"? *Revue Tiers Monde*, *200*, 853–873. https://doi.org/10.3917/rtm.200.0853

Lessard, C. (2007). Les usages politiques de la recherché en education. In V. Dupriez & G. Chapelle (Eds.), *Enseigner* (pp. 69–79). Presses Universitaires de France.

Li Wan Po, A. (1998). *Dictionary of evidence-based medicine*. Radcliffe Medical Press.

Lorenc, T., Pearson, M., Jamal, F., Cooper, C., & Garside, R. (2012). The role of systematic reviews of qualitative evidence in evaluating interventions: A case study. *Research Synthesis Methods*, *3*(1), 1–10. https://doi.org/10.1002:jrsm.1036

Lorence, J. (2006). Retention and academic achievement research revisited form a United States perspective. *International Education Journal*, *7*(5), 731–777. https://files.eric.ed.gov/fulltext/EJ854336.pdf

Matthys, T., & Penxten, J. (2014). *The effect of a structural change of the school holidays on the Belgian economy*. Vlerick & Thomas Cook.

Mercklé, P. (2010). Analyse longitudinale. In S. Paugman (Ed.), *Les 100 mots de la sociologie* (pp. 24–26). Presses Universitaires de France.

Mons, N. (2007). *Les nouvelles politiques éducatives: La France fait-elle les bons choix?* Presses Universitaire de France.

Monseur, C., & Baye, A. (2015). *Les inégalités scolaires d'origines sociale et ethno-culturelle : une possible amplification*. Conseil National d'Évaluation du Système Scolaire. http://hdl.handle.net/2268/193830

Monseur, C., & Lafontaine, D. (2012). Structure des systèmes éducatifs et équité: Un éclairage international. In M. Crahay (Ed.), *L'école peut-elle être juste et efficace?* (pp. 185–220). De Boeck.

Moran, D. J. (2004). The need for evidence-based educational methods. In D. J. Moran & R. W. Malott (Eds.), *Evidence-based educational methods* (pp. 3–7). Elsevier Academic.

Morcke, A. M., Dornan, T., & Eika, B. (2013). Outcome (competency) based education: An exploration of its origins, theoretical basis, and empirical evidence. *Advances in Health Sciences Education*, *18*(4), 851–863. https://doi.org/10.1007:s10459-012-9405-9

Mouraux, D. (1992). *Du rythme d'enfer au rythme d'enfant. Résultats de l'enquête auprès des familles sur le projet d'aménagement des rythmes scolaires*. de Boeck.

Murad, M. H., Asi, N., Alsawas, M., & Alahdab, F. (2016). New evidence pyramid. *Evidence-Based Medicine*, *21*(4), 125–127. https://doi.org/10.1136/ebmed-2016-110401

National Education Commission on Time and Learning. (2005). *Prisoners of time*. NECTA.

Newby, P. (2010). *Research methods for education*. Pearson Education Limited.

Norman, R. (2016). *The changing epistemic governance of European education. The fabrication of the Homo Academicus Europeanus*. Springer.

OECD. (2010). *PISA 2009 results: Overcoming social background. Equity in learning opportunities and outcomes*. PISA, Éditions OCDE. https://doi.org/10.1787/9789264091504-en

OECD. (2011). *Regard sur l'éducation 2011: Panorama*. PISA, Éditions OCDE. https://doi.org/10.1787/eag_highlights-2011-fr

OECD. (2014). *Regard sur l'éducation 2014: Panorama*. PISA, Éditions OCDE. https://doi.org/10.1787/eag_highlights-2014-23-fr

OECD. (2016). *Résultats du PISA 2015 (Volume I): L'excellence et l'équité dans l'éducation*. PISA, Éditions OCDE. https://doi.org/10.1787/9789264267534-f

Pacte pour un Enseignement d'Excellence. (2017). *Socles de savoirs et de compétences du tronc commun: Charte des référentiels*. pactedexcellence.be. http://rire.ctreq.qc.ca/2016/07/enseignement-donnees-probantes/

Pageau, L. (2016, September 13). *Les données probantes et les méta-analyses en éducation*. Réseau d'Information pour la Réussite Éducative. http://rire.ctreq.qc.ca201609donnees-probantes-dt/

Parlement de la Fédération Wallonie-Bruxelles. (2011). *Révisions de la Constitution*. http://www.federation-wallonie-bruxelles.be/index.php?id=109. Pour la dernière ressource, le titre a également changé sur la page Web.

Perrenoud, P. (1996). Lorsque le sage montre la lune . . . l'imbécile regarde le doigt. De la critique du redoublement à la lutte contre l'échec scolaire. *Éduquer et former, théories et pratiques*, *5*, 3–30. https://www.unige.ch/fapse/SSE/teachers/perrenoud/php_main/php_1996/1996_06.html

Petticrew, M. (2015). Time to rethink the systematic review catechism? Moving from "what works" to "what happens." *Systematic Reviews*, *36*(1), 1–6. https://doi.org/10.1186/s13643-015-0027-1

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences, a practical guide*. Blackwell.

Quéré, O. (2017). De la recherche en éducation aux pratiques éducatives. *Revue Internationale d'Education de Sèvres*, *74*, 24–29. https://doi.org/10.4000/ries.5770

Quittre, V., Crépin, F., & Lafontaine, D. (2017). Résultats de l'enquête PISA 2015 en Fédération Wallonie-Bruxelles. *Les Cahiers des Sciences de l'Éducation*, *37*, 1–172. http://www.enseignement.be/index.php?page=26997

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, *73*(2), 417–458. https://doi.org/10.1111/j.1468-0262.2005.00584.x

Rockoff, J. E. (2004). The Impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review*, *94*(2), 247–252. https://doi.org/10.1257/0002828041302244

Sackett, D. L. (1981). How to read clinical journals: I. Why to read them and how to start reading them critically. *Canadian Medical Association Journal*, *124*(5), 555–558. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1705173/

Sackett, D. L. (1989). Rules of evidence and clinical recommendations on the use of antithrombotic agents. *CHEST Journal*, *95*(2), 2S-4S. https://doi.org/10.1378/chest.95.2_Supplement.25

Sackett, D. L. (1995). Surveys of self-reported reading times of consultants in Oxford, Birmingham, Milton-Keynes, Bristol, Leicester, and Glasgow. In W. Rosenberg, W. Richardson, R. Haynes, & D. Sackett (Eds.), *Evidence-based medicine*. Churchill-Livingstone.

Sackett, D. L., Haynes, R. B., & Tugwell, P. (1985). *Clinical epidemiology: A basic science for clinical medicine* (1st ed.). Little, Brown.

Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't. *British Medical Journal*, *312*(7023), 71–72. https://doi.org/10.1136/bmj.312.7023.71

Saussez, F., & Lessard, C. (2009). Entre orthodoxie et pluralisme, les enjeux de l'éducation basée sur la preuve. *Revue française de pédagogie*, *168*, 111–136. https://doi.org/10.4000/rfp.1804

Shaughnessy, J. J., & Zechmeister, E. B. (1990). *Research methods in psychology* (2nd ed.). McGraw-Hill, Inc.

Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Education Researcher*, *31*(7), 15–21. https://doi.org/10.3102/0013189X031007015

Slavin, R. E. (2007). *Education research in an age of accountability*. Pearson Education.

Slavin, R. E. (2016). Getting to scale: Evidence, professionalism, and community. *Journal of Education for Students Placed at Risk*, *21*(1), 60–63. https://doi.org/10.1080/10824669.2014.1110028

Slavin, R. E. (2017). Evidence-based reform in education. *Journal of Education for Students Placed at Risk*, *22*(3), 178–184. https://doi.org/10.1080/10824669.2017.1334560

Slavin, R. E. (2019). How evidence-based reform will transform research and practice in education. *Educational Psychologist*, *0*, 1–13. https://doi.org/10.1080/00461520.2019.1611432

Smith, R., & Rennie, D. (2014). Evidence-based medicine—An oral history. *British Medical Journal*, *348*, 1–3. https://doi.org/10.1136/bmj.g371

Sommer, B., & Sommer, R. (1991). *A pratical guide to behavioral research, tools and techniques*. Oxford University Press.

Testu, F., & Fontaine, R. (2001). *L'enfant et ses rythmes: pourquoi il faut changer l'école?* Calmaan-Levy.

Thomas, G. (2004). Introduction: Evidence and practice. In G. Thomas & R. Pring (Eds.), *Evidence-based practice in education* (pp. 1–18). Open University Press.

Vandenbroeck, M. (2018). L'hégémonie des programmes compensatoires: Un regard international. In P. Ben Soussan & S. Rayna (Eds.), *Le programme "parler bambin": Enjeux et controverses* (pp. 259–278). Erès.

Vandenbroeck, M., Roets, G., & Roose, R. (2012). Why the evidence-based paradigm in early childhood education and care is anything but evident. *European Early Childhood Education Research Journal*, *20*(4), 537–552. https://doi.org/10.1080/1350293W.2012.737238

von der Embse, N. P., Pendergast, L. L., Segool, N., Saeki, E., & Ryan, S. (2016). The Influence of test-based accountability policies on school climate and teacher stress across four states. *Teaching and Teacher Education*, *59*, 492–502. https://doi.org/10.1016/j.tate.2016.07.013

Williamson, J. W., German, P. S., Weiss, R., Skinner, E. A., & Bowes, F. (1989). Health science information management and continuing education of physicians: A survey of U.S. primary care practitioners and their opinion leaders. *Annals of Internal Medicine*, *110*(2), 151–160. https://doi.org/10.7326/0003-4819-110-2-151

Zimerman, A. L. (2013). History of medicine, evidence-based medicine: A short history of a modern medical movement. *American Medical Association Journal of Ethics—Virtual Mentor*, *15*(1), 71–76. https://doi.org/10.1001/virtualmentor.2013.15.1.mhst1-1301

Zorman, M., Bressoux, P., Bianco, M., Lequette, C., Pouget, G., & Pourchet, M. (2015). "Parler": Un dispositif pour prévenir les difficultés scolaires. *Revue Française de Pédagogie*, *193*, 57–76. https://journals.openedition.org/rfp/4890