Forum for Education Studies 2025, 3(3), 3108.
https://doi.org/10.59400/fes3108

Review

Evidence-based education policy in practice: The case of the Netherlands

Geert Driessen

Radboud University, 6525XZ Nijmegen, Netherlands; driessenresearch@gmail.com

CITATION

Driessen G. Evidence-based education policy in practice: The case of the Netherlands. Forum for Education Studies. 2025; 3(3): 3108. https://doi.org/10.59400/fes3108

ARTICLE INFO

Received: 15 April 2025
Revised: 19 May 2025
Accepted: 21 May 2025
Available online: 13 October 2025

Abstract: While evidence-based policy and practice have been on the rise for some time now, this does not apply to education. Evidence-based education (EBE) can be defined as the principle that education practices should be based on rigorously established objective evidence, preferably on randomized controlled trials. In this country-specific case study, the focus is on the Netherlands and, more specifically, developments regarding the educational disadvantage policy. Whereas the Ministry of Education propagates the use of EBE, researchers seem to be rather reluctant. The main research question is how many of the ministry's policy measures have actually been grounded in the principles of evidence-based education? To answer this question, three review studies covering the period 1985–2023 were critically examined. The results can be called shocking: hardly any of the ministry's policy measures have been based upon hard scientific evidence—neither in the past nor in recent years. Explanations for this finding are given, and recommendations are presented.

Keywords: evidence-based education; evidence-informed education; effectiveness; effect size; education disadvantage policy; interventions; elementary education; the Netherlands

1. Introduction

For several decades now, the concept of evidence-based policy and practice has been a well-accepted modus operandi in several domains—at least in theory [1–3]. Medicine is an early and successful example of an evidence-based profession. Since the 1980s, it has considerably gained in prestige by applying proven knowledge and skills based on solid scientific research. The education profession, by contrast, shows signs of dragging its feet [4]. To arrive at an equal footing, educational research would require radical change and would need to invest much more in a high-quality evidence-base [5,6].

1.1. Evidence-based education

Evidence-based education (EBE) can be defined as the principle that education practices should be based on rigorously established objective evidence, preferably on randomized controlled trials (RCT), which are generally viewed as the gold standard of evidence, rather than on tradition, common sense, ideology, personal intuitions, and judgment [7,8]. A comparable definition perceives EBE as a pedagogical approach that aims at improving the quality of the teaching and learning process by applying principles and practices obtained by scientific research [9]. According to Onderwijsraad [10], there are two reasons for the growing attention to evidence-based education. On the one hand, often new methods and approaches are being introduced without any evidence that they are better than the ones already in use. On the other hand, existing knowledge about methods that have been scientifically proven to work finds its way into practice only sparsely.

While RCTs[1], or true experiments, are seen as the highest level of research quality, there are alternatives—though of lesser quality. For instance, the Federal US Every Student Succeeds Act (ESSA)[2] focuses on state and district decision-making, moving from (1) identifying needs to (2) choosing and implementing interventions to (3) examining outcomes [11,12]. ESSA discerns four tiers of evidence; see **Table 1**.

Table 1. The ESSA tiers of evidence (Source: REL Midwest [11]).

	(1) Strong evidence	**(2) Moderate evidence**	**(3) Promising evidence**	**(4) Demonstrates a rationale**
Study design	Well-designed and implemented experimental study	Well-designed and implemented quasi-experimental study	Well-designed and implemented correlational study. Statistical controls for selection bias	Well-defined logic model based on rigorous research
Study results	Statistically significant positive effect	Statistically significant positive effect	Statistically significant positive effect	An effort to study the effects of the intervention is currently underway
Findings related studies	No strong negative findings from (quasi) experimental studies	No strong negative findings from (quasi) experimental studies	No strong negative findings from (quasi) experimental studies	N/A
Sample size & setting	At least 350 participants; more than one setting	At least 350 participants; more than one setting	N/A	N/A
Match	Similar population and setting to your setting	Similar population and setting to your setting	N/A	N/A

As **Table 1** shows, study design is probably the most distinctive and decisive factor of the level of evidence (experimental—quasi-experimental—correlational) attributed to a study. Noteworthy is also the fact that ESSA departs from the criterion of significance. This may be called remarkable, as, for some time now, significance alone has generally been seen as not enough [13]. Effects should be judged in terms of their magnitude as well: how strong is the effect, is it worthwhile, and is it meaningful [14]? Statistical significance refers to the probability that the observed difference between two groups is due to chance. However, the problem is that with large samples, effects will almost always be significant and therefore often meaningless. A (standardized) effect size, like the much-used *d* [15], can then provide a necessary complement; it denotes the difference between the mean outcomes of two different intervention groups (e.g., the experimental and control group in an RCT). A well-known large-scale example of the use of effect sizes in education is the work of Hattie [16], who summarized the results of some 50,000 studies by calculating the effect sizes of teaching strategies to determine which ones function the best. According to Head [17], what lies at the basis of an evidence-based approach are good data, analytical skills, and political support for the use of scientific information.

In addition to the term 'evidence-based' education, 'evidence-informed' education is also employed. Some use these terms interchangeably [18], but others make a clear distinction, not only in terms of level (or hierarchy) but also content [3,19,20]. According to De Wit et al. [21,22], evidence-based education is solely based on rigorous, high-quality, and unbiased scientific research, while evidence-informed education combines various information sources, predominantly:

- Knowledge from (scientific) research;
- Practical knowledge from educational professionals;

- Data from, e.g., tests, student monitoring systems, and observations.

Furthermore, De Wit et al. stress the importance of considering how all the information from the various sources fits in with the specific context and conditions of the (planned) intervention.

1.2. Criticism

While the medical profession nowadays makes use of evidence-based research to the full, this does not apply to the education profession; there it obviously is still a work-in-progress. According to Dekker and Meeter [1], in many countries, educational policymakers have favored EBE programs and interventions for a few decades now. However, this has been accompanied by growing resistance from educational researchers. In the US, the introduction of the Every Student Succeeds Act (ESSA) in 2015 had far-reaching implications because it promoted the use of federal dollars on programs with evidence of effectiveness. ESSA points out that more evidence of what works to increase student success is available, and it therefore encourages the use of strategies with evidence of impact [23]. In the Netherlands, in order to give it a boost, the Ministry of Education's policy has some time ago shifted from encouraging to requiring the use of research in practice [24]. The question is, why is the education profession so reluctant?

Onderwijsraad [10] brings forward the following objections:

- Scientific research can be perceived as a threat to the teaching profession's autonomy;
- The implementation of evidence-based methods requires additional time;
- Scientific research is often limited to a few aspects of children's development;
- Scientific research is expensive, and the deployment of RCTs and CRCTs is complex and demanding.

To this, Dekker and Meester [1] add several other and similar points of criticism:

- Problems with the status of RCTs and generalizations based on them;
- The cost-effectiveness of RCTs;
- The possibility that the focus on experimental research restricts attention to those interventions that can be studied with it;
- Moral implications for the teaching profession; education is not effect-driven but value-driven and therefore an inherently normative profession (cf. Biesta [25]).

In a similar vein, Duijkers [26] mentions a number of risks associated with the introduction of EBE:

- Ignoring contextual and individual differences;
- Limiting teachers' professional autonomy;
- Overemphasis on measurable outcomes;
- Inhibiting innovation and adaptivity in education.

In an explanation, Duijkers argues that the complexity of education cannot be captured in simple, universal solutions. EBE, however, offers the illusion of certainty but often ignores the unique contexts and needs of individual schools, teachers, and students. Therefore, EBE should not be used as a straitjacket but as a source of insights that do justice to the diversity and complexity of student populations. There is not one best way in education; instead, Duijkers advises relying on the teachers'

professionalism, recognizing the importance of context, and embracing a holistic approach that allows for innovation, creativity, and, especially, continued learning on our own. This vision is more or less in line with that of Davies [5] (p. 118), who is of the opinion that EBE 'is not a panacea, a quick fix, cookbook practice, or the provider of ready-made solutions to the demands of modern education. It is a set of principles and practices that can alter the way people think about education, the way they go about educational policy and practice, and the basis upon which they make professional judgments and deploy their expertise'. Onderwijsraad [24] (also see Dekker and Meeter [1]) is of the opinion that it's good to know what works, but also to realize that education is not an exact science, and not everything works everywhere. Therefore, it is important to see under what circumstances, in what context, with what methods, with what target group, and from which scientific (sub)discipline knowledge was collected. Van der Ploeg [20] mentions comparable objections, for instance, that the EBE standard is too radical and does insufficient justice to the relativity of (scientifically acquired) knowledge and that it disregards the complexity of education, the non-deterministic nature of teaching and learning, and the differences among students. Schuler et al. [27] state that the relationship between proven scientific facts and public educational policy is a complex one, which raises questions such as: Whose evidence? Why? For whom and under whose control? Therefore, they recommend promoting knowledge sharing and dialogue between researchers, regulators, and teachers. On the one hand, Gravemeijer and Kirschner [28] focus on problems associated with RCTs. They argue that an evidence-based approach in this context is not feasible, not affordable, and too generalizing. In addition, it shows what works but not how 'it' works, and, moreover, it slows down and conserves. On the other hand, Gravemeijer and Kirschner point out that often ill-considered innovations are introduced that are spread throughout the whole education system far too quickly, namely before there is any evidence that they really work. They also argue that just because more innovative goals are difficult to test does not mean they are not testable at all. Van Woensel [29] relativizes the use of scientific evidence. According to her, such evidence can help in describing how things work and in understanding what is true or false. However, it does not convey nor dictate how to use the evidence. The fact that research shows that a measure is effective does not automatically imply that this measure is an appropriate response to a given policy problem. Decisions to be made about this are reserved for policymakers and their advisors.

From the above, it can be concluded that many of the objections to EBE boil down to an aversion to the (mere) use of quantitative data in general and that of (sophisticated) RCTs in particular. However, Bosker [30] puts these reservations into perspective, especially in relation to the objections brought forward by Gravemeijer and Kirschner [28]. Regarding the objection that RCTs are not feasible, Bosker argues that Gravemeijer and Kirschner assume that students in the control condition receive a placebo, but there is no need for this, as there is a simple solution. Indeed, the control condition can be a regular setting or an existing educational practice. An additional advantage is that this cannot be considered unethical. As to the impossibility of randomization of schools, Bosker is convinced that this need not be a problem if one works with clever solutions, such as schools that are willing to work with the intervention introducing it to a different year group than in the experimental group (the

so-called cross-over design). With regard to the problem of affordability, Bosker points out that huge sums of money go into education with undirected and unproven interventions. If anything is unethical, it is that millions of children are exposed to such interventions (also see Driessen [31]). The point mentioned that it would be too generalizing refers to the problem of early adopters with whom the intervention might work, but not in everyday teaching practice. Bosker argues that it is important to first show that the intervention works (either with the early adopters or with others) and then design a sensible implementation path with initial training and in-service retraining and continuing education of teachers. Dekker and Meeter [1] also analyzed the various objections against EBE and its preference for RCTs. They conclude that adjustments may be in order, but there is no need to abandon EBE. To accomplish this, they recommend three possible directions: (1) study local factors, mechanisms, and implementation fidelity in RCTs; (2) utilize and improve the available longitudinal performance data; and (3) use integrated interventions and outcome measures.

1.3. Research question and method

Policymakers often seem to be convinced of the importance of evidence-based education measures and interventions, but this is not always the case for researchers. In particular, they have objections against the emphasis on the deployment of sophisticated quantitative research methods, especially randomized controlled trials. Though there is a lot of discussion regarding EBE and RCTs, there is hardly any information on the relation between EBE, policy measures, and education practice. This article aims at gaining more insight into the degree to which the Ministry of Education substantiates its own policy measures with the principles of evidence-based education and requires the use of proven interventions when making available (extra) budgets to schools. Or, in other words, how many of the Ministry of Education's policy measures have actually been grounded in the principles of evidence-based education?

This is a country-specific case study, focusing on the Netherlands. Moreover, it concentrates on a specific topic, namely the Educational Disadvantage Policy (EDP). This policy was implemented in the 1970s, and since then billions of euros have been awarded to schools. Because of this longitudinal character, it—to a limited degree—is possible to check whether there are any developments. The selection criteria for this study, therefore, were threefold. In the first place, EDP has always attracted a lot of attention in politics, in practice, and in research (and internationally as well). As a consequence, it could—and still can (though in the US, at the federal level, not any longer)—count on huge (additional) budgets for schools. It has been estimated that since its start, some thirty billion euros have been invested in this policy. It is only obvious that, consequently, one may expect that all this money has been spent efficiently and effectively. This would imply that the money has been used in a targeted manner and without any waste on the one hand, and in such a way that the goals set have been achieved on the other hand. A second reason is a pragmatic one. In the course of years, many hundreds of studies into EDP have been performed, mostly focusing on only one specific aspect (or intervention) at the same time. The resulting knowledge base, as a consequence, is extremely fragmented. Still, there are a few review studies that aim at bringing the outcomes of many single studies together. On

the basis of this, it may be possible to check what the original, underlying policy intentions were, where they were based upon, and what the results were. A third reason pertains to the fact that EDP has been in existence for more than fifty years, which—in principle—makes it possible to check for any developments in the application of evidence-based education. A relevant question then is whether the ministry, since more scientific information and proof has become available, did use this to substantiate its future policies and make adequate choices.

The basis for this state-of-the-art evaluation includes three studies spanning the period 1985–2023. The first is a review study from 2013, which examines to what extent a broad array of policy measures to increase educational equity was based on robust research findings before they were implemented in practice. The second is an overview of interventions that schools had to choose from to receive extra budgets to combat the negative effects of school closures during the COVID-19 pandemic in 2021. The third is a strategic evaluation of the effectiveness and efficiency of policy measures that were taken during the 2017–2023 period and that were meant to increase equality of opportunity in primary education.

2. Evidence-based education in practice

2.1. Educational disadvantage policy interventions

In the Netherlands, the first national policies to prevent and combat educational disadvantage were implemented in 1972. Since then, a wide range of measures and interventions have been developed and applied, aiming at children from low socioeconomic and migrant backgrounds [32]. A rough estimate of the total investments amounts to some thirty billion euros. Strangely enough, in all those years little research has been done to thoroughly investigate their effects. In 2013, the Dutch Parliament asked for a broad, qualitative overview of the policies' effectiveness. A review study was conducted for the years 1985–2012, and a variety of measures were evaluated [33]. Regarding primary and secondary sources, the focus was on empirical-scientific reports and policy documents. In the resulting report, a short description of the policy was given, its background and goals, its implementation, its scientific evaluation, and its outcomes. The most important findings will be summarized here.

2.1.1. The weighting system

From a financial perspective, the most important measure of the Educational Disadvantage Policy (EDP) is the so-called weighting system [34]. On the basis of the social and ethnic background of the children, both elementary and preschools receive (substantial) extra budgets. Relevant is that, because of the Dutch constitutional Freedom of Education, schools are free to spend the additional money in accordance with their own specific needs and preferences, for instance, on class size reduction or extra language tuition for the policy's target group. This measure was not researched before it was introduced in practice, and it has also not been tested appropriately since then. In fact, because of the schools' freedom and the diversity of specific interventions, it is sheerly impossible to evaluate the measure's effectiveness. This means that there was no evidence base for this measure.

2.1.2. Intercultural education

Most EDP measures focus on improving cognitive development. Intercultural education, though, primarily aims at non-cognitive factors, such as contributing to a two-sided process of acculturation of both immigrant and native-Dutch children. This measure was not tested before its implementation either, and in the course of years only a few descriptive studies have been conducted. This implies that there was no evidence base for this measure either.

2.1.3. Immigrant language and culture teaching

This measure was introduced to facilitate immigrant (or so-called guest workers') children's learning of the Dutch language via (firstly) the learning of their so-called mother tongue. In practice, this meant that, for instance, Turkish and Moroccan children received lessons in Turkish and Arabic, alongside lessons in Dutch. Before this measure was implemented, there was only some ambiguous evidence, mainly from the US and the Spanish language. The material used largely came from the children's home countries, and the (immigrant) teachers generally did not have any Dutch (teacher) education. Nothing was tested before this measure was deployed, and only a handful of mostly descriptive and correlational studies have been conducted since then. Therefore, there was no evidence base for this measure.

2.1.4. Dutch as a second language

Right from the beginning, learning the Dutch language has always been a priority for immigrant children. Countless studies have been conducted pertaining to language acquisition, mostly in English-speaking countries, but also numerous in the Netherlands. In addition, lots of materials have been developed, and Dutch as a second language (DSL) was—to some degree—integrated into the teacher training college curriculum. Despite the extensive attention to DSL, nearly all studies were small, descriptive, or correlational; in a methodological sense, they were weak, and their findings were ambiguous. In the course of years, it became clear that a one-sided focus on immigrant children was probably not the best approach, and therefore attention was broadened to all weak students, both immigrant and native-Dutch. Substantive evidence for this measure was lacking.

2.1.5. Pre- and early school education

Since 2000, pre- and early school education (or early childhood education and care; ECEC) has been the main focused EDP intervention. It starts with the premise that it is better to prevent early than to cure later. It aims at preventing target group children from starting formal schooling in arrears and therefore provides compensatory stimulation activities to the children and their parents. The initial interest in this measure was based on a few older experimental programs from the US, which since then have been heavily criticized (cf. Driessen [35]). Later international studies established no unambiguous effects. In the Netherlands, several programs have been developed and implemented, but none of them has been studied or evaluated properly. As far as studies have been done, practically all are descriptive or correlational. A meta-analysis of all Dutch studies found no effect at all. An evidence base did not exist.

2.1.6. Citizenship and social integration

Citizenship education can be seen as a follow-up to the failed intercultural education. According to the Ministry of Education, citizenship education is a mandatory part of the curriculum. However, the legislation is extremely diffuse, and norms about what exactly the goals are simply lacking. In practice, this measure probably does not differ much from intercultural education. Little is known about how schools have implemented this measure, and data on its outcomes and effects are absent. Apart from this, there is no Dutch or international evidence for what exactly works, especially regarding the acquisition of (the 'right') attitudes. All in all, an evidence base for this measure was lacking.

2.1.7. Parental involvement and participation

Parental involvement and participation include a wide variety of active and passive aspects. International studies tend to point to (small) positive effects, but the big problem is that it is often not clear what the measure reported exactly stands for and that all the outcomes of different practical interpretations are lumped together (cf. Driessen [36]). In the Dutch context, hardly any empirical study has been conducted. An evidence base was therefore missing.

2.1.8. The extended school day

The idea is that offering students more learning time will improve their achievement. The number of international studies focusing on this measure is very small, and the outcomes are mostly not positive. In the Netherlands, experiments, typically aiming at non-cognitive activities, were conducted in a few big cities but were not accompanied by adequate evaluations. The results were ambiguous and not based on 'hard' evidence. Despite this disappointing result, the ministry decided to expand this measure. Results of an evaluation showed that there were no significant effects on language and math achievement. An evidence base was thus absent.

2.1.9. Summer schools

Summer schools originate from the US, where children often have three months of summer holidays. When they return to school, their achievement level has often deteriorated, especially that of disadvantaged students. To counter this negative effect of long periods of absence, summer schools were introduced. Evaluative studies do not offer convincing results. In the Netherlands, a few summer schools have opened their doors but have not been evaluated. Thus, there was no evidence base.

2.1.10. Class size reduction

Most elementary schools use the extra budget they receive from the weighting system to reduce the size of their classes. The evidence for this measure is meager and not consistent, however. Often, the STAR project is referred to. This experiment was conducted in the US in the 1980s. It concluded that class size reduction could lead to better language and math achievement. However, this positive effect only appeared when classes were (drastically) reduced from 23 to 15 students. The effect was stronger for disadvantaged students than for middle-class students. More recent meta-analyses did not find such effects, however. Dutch research into class size reduction is scarce; one study showed small differential effects (differing per category of students), and another did not find positive but negative effects. Therefore, for this measure, there

was no evidence base.

2.2. The COVID-19 recovery plan[3]

When, in 2020, the local COVID-19 infections turned into a worldwide pandemic, many countries decided to close their schools to prevent further contaminations. In the Netherlands, the Ministry of Education soon realized that, as a consequence of school closures, student cognitive achievement and well-being would be adversely affected. Therefore, in 2021, the ministry decided to start the Nationaal Programma Onderwijs (NPO; National Program Education; [37]) to help schools to catch up on the COVID-19 arrears, especially those of socioeconomically disadvantaged students. Therefore, it put an unprecedented budget of €8.5 billion at the disposal of the schools. Schools were obliged to dedicate this money (on average €180,000 per elementary school per year, ranging from €5500 to €1.3 million per school) to one or more evidence-based interventions. As the Dutch ministry did not have a database with 'What Works' like scientifically proven interventions (cf. [38,39]) at its disposal, it turned to the British Education Endowment Foundation (EEF; [40]), which was in the process of building such an evidence base. The scientific basis for the interventions' effectiveness is presented on the EEF's website, under Teaching and Learning Toolkit and then specifically the Technical Appendices of each of the interventions. From this base, the ministry took 19 interventions, translated them into Dutch, and—if necessary—adapted them to the Dutch situation and added 3 'own' interventions. The result of these 22 interventions was called the 'NPO menu', and elementary schools had to choose one or more interventions from it. There was a lot of pressure on the schools: they had only a few months to prepare themselves and choose suitable interventions, and they had to spend their extra budget in just two school years[4].

The question here is how 'proven effective' (i.e., evidence-based) the interventions included in the NPO menu really were[5]. To get an idea of this, the evidence provided by EEF was critically examined, thereby focusing on the methodology used in the underlying studies and the outcomes of the interventions, that is, their effects [41]. **Table 2** offers an overview. EEF discerns two types of studies, namely individual studies and meta-analytical studies. A meta-analysis comprises a quantitative summary of a series of individual studies. Effects are expressed in so-called (standardized) effect sizes. To interpret the strength of effect sizes, often the following rule of thumb is applied [15,42]: Less than 0.20 means that there is no or a negligible effect; from 0.20 to 0.50 indicates a small effect, from 0.50 to 0.80 a medium effect, and above 0.80 a strong or very strong effect. A positive effect (+) denotes that the intervention has led to the expected positive results; a negative effect (−) indicates that the intervention has resulted in an undesired negative effect. When there is no effect (0.00), one speaks of a null effect. **Table 2** presents, per intervention, the number of meta-analytical studies and individual studies; the range of the effects (from strong negative to strong positive); the average effect; and notable substantive and methodological comments.

Table 2. Overview of effects of 'proven effective interventions' (effect sizes).

Intervention	Meta-studies/individual studies	Min.–max. effect	Average effect	Comments
Pre- and early-school interventions	12/0	0.15–0.55	0.38	Methodological problems. More recent → weaker effects. Mostly old US studies; UK studies show no consistent effects.
Extending school time	8/5	−0.14–0.40	0.11	Much variation in effects, not consistent. Mostly US studies; little evidence in the UK. Often part of a broader program → origin effect unclear. Problem: absent students.
Summer schools	6/4	0.00–0.43	0.18	Small groups, intensive, experienced teachers → more chance of success. Problem: absent students.
One-to-one tuition	7/7	−0.06–0.70	0.37	Sometimes education of the tutor is important, sometimes not. Substantial variation in effects, not consistent.
Individualized instruction	7/2	−0.07–0.41	0.19	Many old studies. Methodological problems. Much variation in effects.
Small group tuition	4/4	−0.08–1.61	0.31	Often old studies. Mostly secondary education. Not aimed at improving achievement. Enormous variation in effects, not consistent. Quality instruction is probably very important.
Direct instruction*	1 (328)	0.51–0.83	0.33	No direct references. Often old studies in one meta-analysis. Cognitive effects, but no socio-emotional effects.
Peer tutoring	9/4	−0.06–1.05	0.37	A few old studies. Methodological problems. Enormous variation in effects. More recent → stronger effects, often in combination with the use of digital resources.
Feedback	7/0	0.20–0.97	0.63	Almost all (very) old studies. Substantial variation in effects. More recent → weaker effects.
Mastery learning	6/4	0.04–1.64	0.40	Mostly very old studies. Methodological problems. Enormous variation in effects, not consistent. Recent UK studies → weaker effects.
Reading comprehension strategies	8/4	0.10–0.74	0.45	Substantial variation in effects. Mostly US studies. Recent UK studies → weaker effects.
Spoken language interventions	11/7	−0.14–0.91	0.37	Very broad concept with many concrete interpretations. Enormous variation in effects, not consistent.
Well-being interventions*	?	?	?	Very broad concept with many concrete interpretations. No references. Almost all interventions regarding mental health are not effective; effectiveness is unknown.
Sports activities	3/0	0.10–0.80	0.17	Methodological problems. Substantial variation.
Arts participation	5/2	0.03–0.77	0.15	Methodological problems. Substantial variation, not consistent.
Metacognition and self-regulation	11/9	−0.02–0.90	0.54	Effect strongly depends on whether the students can take responsibility for their learning. Methodological problems, enormous variation, not consistent. More recent studies → weaker effects.
Collaborative learning approaches	11/1	0.13–0.91	0.38	Substantial variation, not consistent.
Class size reduction	6/0	0.12–0.34	0.19	Often very old studies from the US. Effects only occur with less than 20 or even 15 students per class.
Teaching assistant interventions	0/15	−0.15–1.50	0.08	Enormous variation, not consistent. Also negative effects.
Facilities and conditions*	?	?	?	In combination with other interventions. No references.
Parental engagement	15/4	−0.14–0.65	0.22	Very broad concept with many concrete interpretations. Substantial variation. Often old studies from the US.
Digital technology	32/2	−0.15–1.13	0.29	Very broad concept with many concrete interpretations. Methodological problems. Enormous variation, not consistent.

* Not in the EEF Toolkit. Source: Education Endowment Foundation [40].

From the data in **Table 2,** it becomes clear that for two interventions it is unknown how effective they are; the average effect of one-third of the interventions is negligible (<0.20); in half of the cases the effect is small (0.20–0.50); and in only two cases the effect is just medium (>0.50). To summarize: insofar as they are effective, nearly all interventions are only weakly effective at most. What is relevant and notable is that an important part of the interventions does not meet the normal standards for effectiveness. Moreover, this assessment is really flattering. The table, indeed, unequivocally shows that there is a huge variation in the strength of the effects; they diverge from negative to strong positive. In most cases, a consistent, unambiguous pattern of effects is lacking. Obviously, the situation is much more complex than assumed. **Table 3** summarizes the most important problematic aspects of this overview of so-called proven effective interventions[6].

Table 3. Overview problematic aspects research into 'proven effective interventions' (trends).

- Interventions include diverse (practical) interpretations; mostly very broad, abstract concepts; often it is not clear what actually is meant.
- No or hardly any directed research; insofar as there is research, it mainly comes from the US and—less—from the UK; Dutch research is mostly completely lacking.
- EEF's conclusions are based on a relatively small number of studies that, in addition, are often (very) dated.
- The (meta)analyses mentioned by EEF show very much overlap, especially when the studies are performed by the same author.
- No studies were performed in the COVID-19 context. The study's goal and target group differ: interventions are mostly developed to prevent, decrease, or eliminate disadvantage of children from lower socioeconomic and ethnic and immigrant backgrounds.
- Studies often do not meet methodological standards.
- EEF reports positive, null, and negative effects, varying from strong positive to strong negative. Consistency is lacking; ambiguity reigns.
- Possible positive effects seem to be over-reported ('cherry-picking'); little attention to the reasons for negative and null effects.
- Possible effects only pertain to specific target groups (e.g., social class, ethnicity, age/grade, primary/secondary school) in combination with specific domains (e.g., language, math, motivation, well-being). In average effect sizes, all this is lumped together, while the implementation of the interventions focusing on specific needs is crucial.
- Interpretation of effects does not always comply with regular standards; 'effective' according to EEF standards, then, in fact, is not effective.
- More recent studies often show weaker effects than older studies.

What can be concluded from the above? Firstly, it is remarkable that a Ministry of Education has no overview of what really works in schools, what are scientifically proven interventions, and what are not. Secondly, it comes as a surprise that this ministry then falls back on a foreign evidence-base and prescribes its interventions to all eligible Dutch schools without checking whether these interventions really are evidence-based. What is positive is that the ministry is now conducting a large longitudinal study to monitor whether the NPO menu has the desired effect [37]. This evaluative study also aims at finding out which of (a limited number of) the interventions is effective. A big problem, however, is that, in the course of years, schools have had a lot of freedom in adapting and modifying interventions, and that interventions may differ per student, and students may have participated in various interventions at the same time. Though school boards should motivate their NPO expenses in their annual report, according to Algemene Rekenkamer [43], this will not shed any light on the effectiveness of the interventions, as the link between extra budget, interventions, and results is missing.

2.3. Policy measures to improve equal educational opportunity[7]

Since early in the 1970s, equality of educational opportunity has been high on the Dutch political agenda, and since then, dozens of billions of euros have been invested

[32]. Oddly enough, there have been hardly any adequate policy evaluations. This is certainly notable, as the government and each of its ministries proclaim that it is responsible for spending public money in a useful and relevant, economical and efficient, and careful and effective way [44]. The main question, therefore, is whether all these investments have resulted in achieving the policy's goal of bridging the gap between the children of the rich and the poor. Numerous recent studies have shown that this goal has not been achieved; on the contrary, many studies show that the gap is widening, both in the Netherlands and internationally [45,46]. Recently, Reuvers et al. [47] performed a strategic evaluation in order to get insight into the effectiveness and efficiency of all the policy measures from the 2017–2023 period that were meant to increase equality of opportunity in primary education. The focus was on the extent to which the policy theory proved to be correct and the intended output and outcomes were realized. A total of 44 measures (1 had 2 goals) were thus reviewed by Reuvers et al. [47] using a so-called effects ladder, with scores ranging from 0 (implicit) to 4 (effective/not effective). In **Table 4** the results are summarized. The budget mentioned there concerns the year 2023; the total budget of the 7 years studied amounted to €6,879,814,000.

Table 4. Policy measures to improve equal educational opportunity, 2017–2023 (Source: Reuvers et al. [47]; own calculations).

Level	Quality	Description	Number	%	% No study	Budget 2023 (×€1000)
0	Implicit	The measure is not or is badly described; or it is not clear for what problem the measure offers a solution	2	4%	50%	€91
1	Measure and goal defined	It is not well described what the measure is/does, and it is not clear for what problem the measure offers a solution	7	15%	86%	€24,400
2	In theory effective	The measure is well substantiated, and it is clear what mechanism or policy theory the measure departs from, and possible negative side effects are taken into account	16	36%	56%	€429,415
3	Plausibly effective/plausibly not effective	In actual practice there are good/bad experiences, or the measure's intended output has/has not been realized, and it is likely/not likely that this output will contribute to the intended outcome in practice on the basis of international examples, scientific literature, or research about the form and organization of policy	5/14	42%	11%	€1,017,410 / €7250
4	Effective/not effective	Evaluation research is available that confirms/invalidates a relationship between the measure's output and the intended outcome	0/0	0%	0%	€0/€0
?		Not clear	1	2%	100%	
Total			**45**	**100%**	**42%**	**€1,478,566**

What **Table 4** shows is actually a rather shocking and disconcerting picture. None of the measures has a score of 4; in other words, there is no measure whatsoever of which it has scientifically been established that it is effective. What is more, 42 percent of the measures have not been researched at all, while the total budget amounted to nearly €7 billion! The table also shows that 42 percent of the measures are likely to be effective or not effective (score 3). However, that judgment is not based on adequate effect studies but on impressions obtained on the basis of a mishmash of sources, such as descriptive research, various monitors, and research not meant for effect evaluations. For that matter, Reuvers et al. [47] are rather docile regarding their

qualification of the quality of the studies reviewed (as far as studies were carried out at all). They conclude that of none of the measures is it possible to determine their effectiveness with certainty. In addition, they note that it is sometimes also difficult to receive an indication of it. According to them, to arrive at this, research methods are needed that prove causality, for instance, an experiment with a control group or a study with a quasi-experimental design. The first type is not available at all; the second type is only scarcely applied. Reuvers et al. [47] are of the opinion that it is not strange that only a few of the measures have been researched to determine their effectiveness. According to them, this type of research is often complicated and expensive, and, moreover, it can lead to ethical questions when the use of control groups results in unequal treatment. That such research is complex is certainly true, but to therefore systematically refrain from such evaluations would be a rather bizarre motive. In addition, the supposed high costs of such studies are relative and point to short-term thinking. What do a few million research expenses mean in the light of a yearly budget of €1.5 billion for measures of which it is completely unknown whether they are effective or worse, have an adverse effect? Even more important is to think about how much can be saved once it is clear what works and what doesn't work. The ethical objection to such research is even more ridiculous. For many decades now, millions of children have been exposed to measures and interventions of which one has no idea if they are effective or even have a negative effect. To give a concrete example, let's take class size reduction. Elementary schools receive huge supplementary budgets; during the COVID-19 period, they even received three budgets, namely the regular lump sum, the (extra) EDP money, and the (extra) COVID-19 money. Most of the extra money was used to create smaller classes. From the above inventory, it can be concluded that there is no hard scientific evidence that class size reduction helps close the achievement gap between low- and high-socioeconomic status children[8]. Recent research even points to negative effects; that is, children in smaller classes achieve less. The latter means that, had these children been in 'normal' classes, they would probably have achieved the same or even better. From this perspective, one could argue that during all those years, billions of euros have been wasted, which could have been invested in interventions that really are evidence-based. Another example pertains to pre- and early school education. In the Netherlands, a number of such programs have been 'recognized' by the ministry, though there is no evidence that they work, that is, prevent young disadvantaged children from starting elementary school with arrears. Still, this intervention has been heavily subsidized for several decades now. It has been suggested more often to design a large-scale experiment to prove that such interventions work (or not). An argument that is usually raised, however, is that it would be unethical (or unfair) to expose one group of children to an experimental condition (who then may achieve better) and another group not (the control group in the business-as-usual condition). What opponents of an experiment forget is that the billions could have been spent better (i.e., more effectively), and that all the children in the control group may now be exposed to the better (or worse) business-as-usual condition—no one knows. How ethical is all this (cf. Bosker [30])?

3. Conclusion and discussion

This country case study focused on the implementation of evidence-based measures and interventions regarding educational disadvantage policies in the Netherlands. Three studies covering the period 1985–2023 were critically reviewed. The main question was, how many of the Ministry of Education's policy measures have actually been grounded in the principles of evidence-based education?

The review of the first period, 1985–2012, showed that, insofar as research has played a role, it nearly all originated from other countries, specifically the US, and to a much lesser extent, the UK. Only a limited number of studies were conducted in the Netherlands. It concerned descriptive and correlational studies and hardly any ('the gold rule') experimental studies. Moreover, the results of these studies were typically ambiguous and not consistent; effects were generally small, there were many null effects, and even negative effects. Of course, in the first decades of this period, EBE had hardly penetrated into the Ministry of Education, but several of the measures still apply today, while there is as yet no scientifically sound evidence that they are effective.

The results of the COVID-19 period (2020–2022) were actually shocking. While, initially, the ministry required the schools to spend the huge extra budgets they received on evidence-based interventions, it turned out that the interventions they had to choose from were not at all as evidence-based as was suggested. Furthermore, soon the ministry lowered its requirements, and—as a result—each school was doing what it pleased.

The third period, 2017–2023, is undoubtedly the most interesting, as at that time the discussion around the effectiveness of measures and interventions played an important role in research and policy. The results showed that there is no measure whatsoever of which it has been scientifically established that it is effective. Even more surprising is that nearly half of the measures have not been evaluated at all.

Answering the main question of this article is thus rather simple (and, in fact, really disappointing): hardly any of the policy measures have been based upon hard scientific evidence. The question now is, how is this possible, and why is it still happening? Ledoux et al. [48] explain this in part by the circumstance that the rapid dynamics of politics are at odds with the often long lead time of scientific research. Ministers of Education are only in office for a restricted period of time and have no patience for longitudinal studies (that are inherently needed for proving that something works). Moreover, they often have no background in education and certainly not in research. Faasse et al. [49] come up with some more general bottlenecks:

- Issues of the day offer little room for reflection;
- No overview of what knowledge already exists, an overwhelming amount of information, and a lack of understanding of what is going on in practice;
- Accountability may get in the way of learning;
- Publicly asking politically sensitive questions may be difficult for departments;
- Personnel policy focuses on advancement and political-strategic skills, which makes building substantive knowledge difficult.

Advisory Council for Science, Technology and Innovation (AWTI; [50]) notes partly similar bottlenecks:

- Researchers and policymakers do not get in touch with each other enough;
- Knowledge remains with individuals, who then often rotate, causing knowledge to disappear;
- There is a lot of process expertise but no substantive expertise among government policymakers;
- There is little contact between policymakers and implementing organizations, so the former does not know how the policy actually works out;
- Lack of time.

In order to bring policy and research closer together, Waterreus and Sipma [51] recommend that, especially at the stages of policy preparation and policy formulation, it is important to search broadly for solutions that have been proven to be effective in order to increase the likelihood that the policy goal will actually be achieved with the policies put in place. To help accomplish this, they have a few suggestions, such as investing in knowledge management, including dissemination of research findings; organizing interaction between policymakers and researchers; paying attention to education policy research; and stimulating a knowledge-oriented HRM culture. In their 'Manifesto for evidence-based education', Coe and Kime [18] (p. 8-9) also provide (partly overlapping) suggestions, such as strengthening the working relationships between practitioners, researchers, funders, and policymakers. A shift is needed that 'must see the production of research evidence in education as problem-oriented, driven by demonstrable need, and in collaboration with school- and college-based practitioners'. What is also needed is strengthening 'the safeguards that allow policymakers to do the difficult but right.' This includes 'bodies that are independent of government whose role it is to provide distance between political incentives for short-term, simplistic solutions and what evidence supports as best bets'.

- More specifically for the Dutch situation, some concrete recommendations for the Ministry of Education can be formulated: Attract more policymakers with a background in research and methodology, in combination with deep domain expertise (in contrast to the present practice of ministry officials having to rotate every few years).
- Organize more interaction between policymakers, implementers, and researchers, for instance, a few times per year, structural meetings to discuss and prepare new policy initiatives, their backgrounds, goals, contents, scientific foundation, feasibility, implementation, and evaluation. Such meetings would also be very useful for government and parliament members, preferably when plans are still at an early stage of development.
- Create and then update a topical database of 'What works in education' filled with (national and international) high-quality studies, and with much attention to relevant details and specifics, especially methodology, effects, and conditions (instead of what nowadays is rather common; see, e.g., EEF).
- Realize that obtaining true evidence costs time and money (and will often exceed the parliamentary term; there are no quick fixes).
- Finance true experiments with a large-scale and longitudinal design (and be critical of low-quality studies and ambiguous evidence).

Though such suggestions certainly may be helpful, from the above it has become

clear that, indeed, such changes will take quite some time and the commitment and involvement of all stakeholders.

Institutional review board statement: Not applicable.

Informed consent statement: Not applicable.

Conflict of interest: The author declares no conflict of interest.

Notes

[1] In addition to RCTs, so-called Cluster Randomized Controlled Trials (CRCT) are applied, where bigger experimental and control groups are involved to neutralize possible effects of random differences in group composition [10,30].

[2] The Every Student Succeeds Act (ESSA) is a law that governs the United States K-12 public education policy and which, in 2015, replaced the No Child Left Behind (NCLB) Act.

[3] This section is partly based on Driessen [31].

[4] Initially, the ministry demanded that the schools implement evidence-based interventions. Later on, the ministry became more lenient and also allowed the schools to choose from 'promising' interventions [21]. Instead of speaking of 'evidence-based', the ministry then used the qualification 'evidence-informed' [52].

[5] It should be mentioned that the EEF toolkit is a work-in-progress and is being updated periodically.

[6] Just one (randomly chosen) example, extending school time (EST), to show how outdated and imprecise the studies used by EEF [53]. EEF provides the evidence for the effects of applying (diverse) forms of EST on reading outcomes in the primary phase. A total of 43 publications are listed, dating from 1974 to 2016. It is important to realize that the empirical data for these studies were often collected years earlier. No less than 32 publications date from 2000 or earlier; that is, three-quarters of the evidence dates from studies that are at least 25 years old. The most recent publication is, again, ten years old. EST comprises an amalgam of interpretations, ranging from year-round schooling to after-school programs and kindergarten programs. The effect measures vary from (national) standardized test scores to locally developed instruments. Nearly all the studies cited were conducted in (different) American states. Researchers should truly question what the implications of all this are for the reliability and validity of the findings and possible support to policy-making.

[7] This section is partly based on Driessen [41].

[8] Though it may help reduce the workload of teachers, especially those at schools with many disadvantaged children.

References

1. Dekker I, Meeter M. Evidence-based education: Objections and future directions. Frontiers in Education. 2022; 7: 941410. doi: 10.3389/feduc.2022.941410
2. European Commission. Support Mechanisms for Evideince-Based Policy-Making in Education. Publications Office of the European Union; 2017. doi: 10.2797/575942
3. Jepma IJ, Willemsen M, Haagsman A, et al. Kennisgedreven Onderwijs. Onderzoek naar Evidence-Informed Werken in het Funderend Onderwijs. Available online: https://sardes.nl/publicatie/kennisgedreven-onderwijs-onderzoek-naar-evidence-informed-werken-in-het-funderend-onderwijs/ (accessed on 1 April 2025).
4. Voyager Sopris Learning. What are evidence-based practices in education? Available online: https://www.voyagersopris.com/vsl/blog/what-are-evidence-based-practices-in-education (accessed on 2 April 2025).
5. Davies P. What is evidence-based education? British Journal of Educational Studies. 1999; 47(2): 108-121. doi: 10.1111/1467-8527.00106
6. Hargreaves D. Teaching as a Research-Based Profession: Possibilities and Prospects. Teacher Training Agency; 1996.
7. Wikipedia. Evidence-based education. Available online: https://en.wikipedia.org/wiki/Evidence-based_education (accessed on 7 March 2025).
8. Wikipedia. Evidence-based policy. Available online: https://en.wikipedia.org/wiki/Evidence-based_policy (accessed on 8 March 2025).

9. ProFutoro. Evidence-based education: Substantiated innovation. Available online: https://profuturo.education/en/observatory/trends/educacion-basada-en-evidencias-innovando-con-fundamento/ (accessed on 7 March 2025).
10. Onderwijsraad. Naar Meer Evidence Based Onderwijs. Onderwijsraad; 2006.
11. REL Midwest. ESSA Tiers of evidence. What you need to know. Available online: https://ies.ed.gov/ies/2025/01/essa-tiers-evidence (accessed on 5 April 2025).
12. Evidence for ESSA. Standards and procedures Version 2.2 February, 2024. Available online: https://www.evidenceforessa.org/wp-content/uploads/2024/02/FINAL-Standards-and-Proc-02-12-24.pdf (accessed on 1 April 2025).
13. Sullivan G, Feinn R. Using effect size—or why thepvalue is not enough.Journal of Graduate Medical Education. 2012; 4(3): 279-282. doi: 10.4300/JGME-D-12-00156.1
14. Coe R. It's the effect size, stupid: What 'effect size' is and why it is important. In: Proceedings of the 2002 Annual Conference of the British Educational Research Association; 12-14 September 2002; University of Exeter, Exeter, England.
15. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Available online: https://www.utstat.toronto.edu/brunner/oldclass/378f16/readings/CohenPower.pdf (accessed on 7 April 2025).
16. Hattie J. Visible Learning: A Synthesis of over 800 Meta-Analyses Relating to Achievement. Routledge; 2008.
17. Head B. Evidence-based policy: Principles and requirements. In: Productivity Commission, editors. Strengthening Evidence Based Policy in the Australian Federation: Roundtable Proceedings. Productivity Commission; 2010. pp. 13-26. Available online: https://www.pc.gov.au/research/supporting/strengthening-evidence/roundtable-proceedings-volume1.pdf#page=21 (accessed on 1 April 2025).
18. Coe R, Kime S. A (New) Manifesto for Evidence-Based Education: Twenty Years on. Available online: https://evidencebased.education/new-manifesto-evidence-based-education/ (accessed on 7 April 2025).
19. Bennett T. Evidence-based education is dead—long live evidence-informed education: Thoughts on Dylan Wiliam. Available online: https://behaviourguru.blogspot.com/2017/11/evidence-based-education-is-dead-long.html (accessed on 7 April 2025).
20. Van der Ploeg P. Nonsense-based onderwijs en zelf-diskwalificatie. Geïllustreerd aan het Proces Communicatie. Rijksuniversiteit Groningen; 2019.
21. De Wit F, Dubbeld A, Hoff V, et al. Evidence-Informed Werken met Open Leermateriaal. Available online: https://www.openleermateriaal.nl/app/uploads/2024/09/Visiedocument-Evidence-informed-werken-IOL.pdf (accessed on 7 April 2025).
22. MinOCW. Evidence-informed werken. Available online: https://www.masterplanbasisvaardigheden.nl/aan-de-slag-met-de-basisvaardigheden/evidence-informed-werken (accessed on 10 March 2025).
23. Evidence for ESSA. About. Available online: https://www.evidenceforessa.org/about/ (accessed on 9 April 2025).
24. Onderwijsraad. Denk mee over het gebruik van evidence in onderwijspraktijk en onderwijsbeleid. Available online: https://www.onderwijsraad.nl/actueel/nieuws/2025/03/24/denkt-u-mee-evidence (accessed on 5 April 2025).
25. Biesta G. Why 'what works' won't work: Evidence-based practice and the democratic deficit in educational research. Educational Theory. 2007; 57(1): 1-22. doi: 10.1111/j.1741-5446.2006.00241.x
26. Duijkers A. De Keerzijde van 'Evidence-Based' Onderwijs: Vier Grote Risico's. Available online: https://www.remindlearning.nl/de-keerzijde-van-evidence-based-onderwijs (accessed on 7 April 2025).
27. Schuller T, Jochems W, Moos L, et al. Evidence and Policy Research.European Educational Research Journal. 2006; 5(1): 57-70. doi: 10.2304/eerj.2006.5.1.57
28. Gravemeijer K, Kirschner P. Naar meer evidence-based onderwijs? Pedagogische Studiën. 2007; 84(6): 463-472.
29. Van Woensel L. Evidence for Policy-Making. Foresight-Based Scientific Advice. European Parliamentary Research Service; 2024.
30. Bosker R. Naar meer evidence based onderwijs! Pedagogische Studiën. 2008; 85(1): 49-51.
31. Driessen G. Driessen duikt in NPO. Didactief Online. 2021. doi.org/10.5281/zenodo.10902957
32. Driessen G. The Many Facets of Educational Disadvantage. Policies, Interventions, Effects. Eliva Press; 2022. doi: 10.5281/zenodo.6641161
33. Driessen G. De bestrijding van onderwijsachterstanden Een review van opbrengsten en effectieve aanpakken. ITS; 2013. doi: 10.5281/zenodo.10825576

34. Ladd H, Fiske E. The Dutch Experience with Weighted Pupil Funding: Some Lessons for the U.S. Duke University; 2009.
35. Driessen G. The fragile foundation of pre- and early-school programs for disadvantaged children. Forum for Education Studies. 2024; 3(1): 1-16. doi: 10.59400/fes1869
36. Driessen G. Parental involvement: Types and effects. In: Annual of Sofia University 'St. Kliment Ohridski'. Kliment Ohridski University Press; 2021. pp. 7-30. doi: 10.5281/zenodo.6760016
37. MinOCW. Plan van aanpak Monitoring en evaluatie Nationaal Programma Onderwijs.
a) Available online: https://www.nponderwijs.nl/documenten/publicaties/2021/06/25/pva-monitoring-en-evaluatie-np-onderwijs (accessed on 10 June 2021).
38. Institute of Education Sciences. What Works Clearinghouse. Available online: https://ies.ed.gov/ncee/wwc/ (accessed on 29 March 2025).
39. Proven Tutoring. The nation's only clearinghouse for research-proven tutoring models. Available online: https://proventutoring.org/ (accessed on 7 March 2025).
40. Education Endowment Foundation. Teaching an d Learning Toolkit. An accessible summary of education evidence. Available online: https://educationendowmentfoundation.org.uk/education-evidence/teaching-learning-toolkit (accessed on 1 June 2021).
41. Driessen G. Het ministerie doet maar wat. Didactief Online. 2025. doi: 10.5281/zenodo.14794035
42. Nederlands Jeugdinstituut. Effectgrootte. Available online: https://www.nji.nl/effectieve-jeugdhulp/effectgrootte (accessed on 11 March 2025).
43. Algemene Rekenkamer. Verantwoordingsonderzoek. Ministerie van Onderwijs, Cultuur en Wetenschap. Algemene Rekenkamer; 2022. Available online: https://www.rekenkamer.nl/onderwerpen/verantwoordingsonderzoek/documenten/rapporten/2022/05/18/resultaten-verantwoordingsonderzoek-2021-ministerie-van-onderwijs-cultuur-en-wetenschap (accessed on 22 January 2025).
44. Algemene Rekenkamer. Bestrijding Onderwijsachterstanden. Available online: https://www.rekenkamer.nl/publicaties/rapporten/2001/01/30/bestrijding-van-onderwijsachterstanden (accessed on 22 January 2025).
45. De Wolf I. Kwaliteit basisonderwijs staat onder druk. ESB. 2023; 108(4819): 104-108.
46. Hanushek E, Peterson P, Talpey L, et al. The achievement gap fails to close: Half century of testing shows persistent divide between haves and have-nots. Education Next. 2019; 9(3): 8-17.
47. Reuvers L, Geurts A, Bieleman L, et al. Kiezen voor Kansen. Strategische Evaluatie Kansengelijkheid in het Funderend Onderwijs 2017-2023. Andersson Elffers Felix; 2024. Available online: https://www.poraad.nl/system/files/2024-11/Kiezen-voor-kansen-strategische-evaluatie-kansengelijkheid-in-het-funderend-onderwijs-2017-2023.pdf (accessed on 7 March 2025).
48. Ledoux G, Van Eck E, Heemskerk I, et al. Impact van de Commissie Dijsselbloem op Onderwijsbeleid. Kohnstamm Instituut; 2014.
49. Faasse P, Van Elzakker I, Diederen P. Kennis, Kunde, Beleidskeuzes. Rathenau Instituut; 2020.
50. AWTI. Rijk aan Kennis. Naar een Herwaardering van Kennis en Expertise in Beleid en Politiek. Available online: https://www.awti.nl/documenten/adviezen/2021/02/17/advies-rijk-aan-kennis (accessed on 22 January 2025).
51. Waterreus I, Sipma L. Kennisingrediënten voor Onderwijsbeleid. Verkenning naar de Mogelijkheden om de Impact van Onderzoek op Onderwijsbeleid te Versterken en de Rol van het NRO daarbij. NRO; 2013.
52. MinOCW. Nationaal Programma Onderwijs. Zesde Voortgangsrapportage. Ministerie van Onderwijs, Cultuur en Wetenschap; 2024.
53. Education Endowment Foundation. Extending school time. Available online: https://educationendowmentfoundation.org.uk/education-evidence/teaching-learning-toolkit/extending-school-time/technical-appendix (accessed on 15 May 2025).