Research on Preschool and Primary Education
https://ojs.luminescience.cn/RPPE

LUMINESCIENCE
https://luminescience.cn

Review Article

Evidence-based educational interventions as the solution to learning gaps caused by the COVID-19 pandemic?

Geert Driessen*

Radboud University, Nijmegen, The Netherlands (retired)

*Corresponding to: Geert Driessen, Email: driessenresearch@gmail.com

Abstract: The COVID-19 pandemic had a significant impact on the cognitive and non-cognitive development of students in education. In the Netherlands, a plan was developed early to counteract the anticipated negative effects. Drawing on an in-depth literature review, this article examines how the plan was structured, implemented, and evaluated. The total budget for the plan was an unprecedented 8.5 billion euros. The core of the plan compromised a list of evidence-based interventions from which schools could select. The central question is to what extent it can be demonstrated that the plan and its interventions resulted in the same level of cognitive and non-cognitive development after the plan was completed as before the pandemic began. The results show that this goal was partially achieved, but a causal link between the plan and outcomes cannot be demonstrated. A comparison with other Western European countries indicates that this conclusion applies to all these countries.

Keywords: COVID-19, Learning gaps, Disadvantaged students, Evidence-based education, Meta-analyses, Effect sizes, Toolkits

1. Introduction

Soon after the COVID-19 epidemic broke out in 2020 and schools had to close their doors for longer periods of time, the education sector realized that this could have far-reaching consequences for its students, and in particular for disadvantaged students (Di Pietro, 2023; EEF, 2022; Reimers & Schleicher, 2020; Volente & Klinger, 2026). The focus of the present country-specific case study is on how the Ministry of Education in the Netherlands dealt with that and what the outcomes were of this exceptional and unprecedented endeavor. Special attention is being paid to methodological and practical aspects of the choices made. More specifically, the research questions are as follows:

How has the Dutch Ministry of Education addressed the expected educational disadvantages resulting from COVID-19?

- Which measures and interventions were central to this?
- To what extent were these interventions theoretically and empirically substantiated?
- To what extent have the stated objectives been achieved?
- Are there differences between the approach and results in the Netherlands and other Western European countries?

This study combines policy analysis with subsequent methodological critique. It is based on a highly varied literature review, as well as analyses and calculations. The input for this consists of a range of official policy documents, their elaborations, critiques from official advisory

Received: Jan.29, 2026; Revised: Mar.20, 2026; Accepted: Apr.23, 2026; Published: Apr.30, 2026

DOI: https://doi.org/10.55976/dma.42026152657-74

bodies and individual researchers, databases accessed, international scientific articles and books, monitoring reports, and impact and effect studies. All this information is drawn from the author's extensive historical archive on these topics and further obtained via the snowball method. The central keywords were: COVID-19; educational disadvantage; and evidence-based interventions.

Below, we first discuss the COVID-19 master plan developed by the Ministry. The methodological underpinnings will receive considerable attention. We will also address the extensive criticism the plan received. Next, we will examine the extent to which the objectives have been achieved. A summary of approaches adopted in other Western European countries will then be presented. Finally, conclusions will be drawn and recommendations made.

2. The National Education Plan

In early 2021, the Dutch Ministry of Education in a remarkably short period to cope with the expected negative consequences of the epidemic, the National Education Plan (NEP; "Nationaal Plan Onderwijs"). The NEP aimed to help students catch up on learning delays and delays in social-emotional development and well-being caused by the COVID-19 epidemic. (In primary education, school closures lasted 11.5 weeks, while in secondary education they lasted up to 16 weeks; François & De Witte, 2025.)

According to this plan, the education sector would receive an unprecedented extra budget of totaling 8.5 billion euros; of this, 5.8 billion euros were allocated to the primary and secondary education phase (MinOCW, 2021a, 2021b, 2021c). For example, in the school year 2022/2023, primary schools annually received €500 per student and secondary schools €1,000 per student. On average, primary schools (with an average of 225 students) received €112,500, and secondary schools (on average 1,450 students) €1,450,000 (Heijsters et al, 2020). However, this only concerned the basic extra budget. Because it was expected that socio-economically disadvantaged students would have greater delays and would therefore need more help, schools visited by such students received even more in addition to the basic budget, up to €1,406,560 extra (MinOCW, 2022). Initially, schools were required to spend these huge sums within two years, i.e., in the 2021/2022 and 2022/2023 school years. Any money not spent in this period was to be returned to the Ministry. Later, it was decided that the money could be spent for another two years, up to and including the 2024/2025 school year.

The Ministry was in a great hurry; after all, it considered the problems to be very urgent. The NEP was announced in a letter to schools published in March 2021. The schools were then required to conduct a School Scan in April, which identified their specific problems and needs. Next, they had to formulate their goals and develop a School Program for addressing these problems. This program had to be based on an evidence-based approach. In the April – June period, schools therefore had to choose from a so-called Menu Card containing a list of (initially) 22 scientifically proven effective interventions. (Interventions are also called "strands" or "approaches". For readability reasons, the term "interventions" will be used throughout the rest of this article.) (In fact, the Ministry had also compiled a list of interventions that have been proven to be ineffective; research had shown that they have no effect on learning gains. Schools were forbidden to use these interventions with the funds allocated to the NEP; MinOCW, 2021c.) The school's Participation Council had the right to approve the program and assess whether the procedure had been followed correctly and whether everyone in the school was sufficiently involved. After approval (June), the schools had to start implementing the program immediately following the summer holidays (August). The school boards were expected to account for the implementation and results of the school program and the expenditure of resources yearly through the annual report. The implementation was monitored and evaluated by the Inspectorate of Education and several independent research agencies, which would result in a final report at the end of 2025 (MinOCW, 2021c). In the following, the focus is on the core of the NEP, its Menu Card and evidence-based interventions, more specifically, on its backgrounds, implementation, and evaluation. The main question is whether the NEP has ultimately resulted in its intended effects, which was catching up on learning and developmental delays. And if not, why not?

3. The Menu Card

3.1 Evidence-based interventions

Evidence-based education (EBE) has been on the rise for some decades now (Dekker & Meeter, 2026; Gorard et al., 2020). In The Netherlands, just like in several other countries, the Ministry promotes, and will soon move to the obligation, to work in an evidence-based (or evidence-informed) manner (Peeters, 2025; Tweede Kamer, 2023). There are several definitions of EBE. Generally, it is viewed as the principle that education practices should be based on rigorously established objective evidence.

Because evidence-based education was already on the political agenda, the Ministry aligned itself with this approach for the NEP. However, as there was no overview of evidence-based interventions available in the Netherlands, the Ministry turned to the UK, where, since 2011, the Education Endowment Foundation (EEF) has aimed to break the link between family income and educational achievement by helping schools use evidence to improve teaching and learning, especially for disadvantaged students. The EEF achieves this by funding research, summarizing global evidence, and creating practical tools. The EEF is a government-designated What Works Centre for improving education outcomes for school-aged children (https://educationendowmentfoundation.org.

uk/). The core products of the EEF includes two toolkits, the Teaching and Learning Toolkit and the Early Years Toolkit. These are designed to support teachers and school leaders in making decisions about how to improve learning outcomes, particularly for disadvantaged children. Seventy percent of school leaders in the UK use the toolkits to inform their teaching and learning decisions. The toolkits present over 40 interventions to improve teaching and learning. The evidence for their effectiveness is based on a thorough analysis of more than 3,000 individual studies and meta-analyses, and summarized in terms of: (1) its average impact on attainment; (2) the strength of the evidence supporting it; and (3) its cost (EEF, 2023).

A group of Dutch experts, in collaboration with the EEF, selected several interventions they considered appropriate for the task at hand. These were translated into Dutch. As the toolkits focused on cognitive achievement, a few interventions targeting well-being and socio-emotional development were added. Expert assistance was provided by, among others, the Netherlands Youth Institute (https://www.nji.nl/english). This, initially, resulted in a total of 22 interventions. In a later phase, 2 more interventions were added, namely Writing instruction and Phonemic awareness (MinOCW, 2022).

3.2 Effect sizes

An overview of the 22 interventions is presented in Table 1. The coefficients in this table refer to effect sizes (ES), which indicate the intervention's impact on cognitive and non-cognitive attainment. The best-known index for ES is Cohen's d, which can be either negative or positive. A positive value indicates a favorable effect of the intervention; a negative value indicates a counterproductive effect. When there is no effect (d = 0.00), it is called a zero or null effect. Theoretically, this index can take any value, but d typically fluctuates between -2.00 and +2.00.

There is considerable discussion regarding the interpretation of ESs. Cohen (1988), the "inventor" of ES d, provides the following rule of thumb for interpreting the strength of an ES: 0.20 = small (trivial); 0.50 = medium (noticeable); and 0.80 = large (substantial). According to Hattie (2009), famous for his synthesis of over 800 meta-analyses (a so-called meta-meta-analysis) of educational intervention effects on achievement, ESs should be interpreted as follows: 0.20 = small; 0.40 = medium; and 0.60 = large. The average ES across all his meta-analyses is 0.40, which he considers the typical effect of all possible influences in education and recommends as a benchmark or hinge-point to judge effects in education. ESs above 0.40 are, in his view, within the Zone of Desired Effects. EEF (2023) uses a different classification: -0.04 – 0.04 = no or very low impact; 0.05 – 0.18: low impact; 0.19 – 0.44 moderate impact; 0.45 – 0.69 high impact; 0.70 – 1.00 very high impact. Relevant is that EEF focuses on experimental and quasi-experimental research, and quantitative educational and cognitive achievement, and excludes attitudinal, affective or motivational outcomes (Higgens et al., 2024). Kraft (2020) applies entirely different criteria: <0.05 = small; 0.05 – <0.20 = moderate; ≥0.20 = large. It should be mentioned that this benchmark is restricted to a very specific sort of study, namely causal research that evaluates effects of educational interventions on standardized student achievement. According to Kraft, raising educational achievement is difficult. Some 25 percent of the ESs from RCTs (Randomized Controlled Trials – the "gold standard" in research) of education interventions with standardized test scores are negative or zero; moreover, many more small positive effects cannot be distinguished from zero. The Netherlands Youth Institute (NJi, 2026), which was involved in the selection of the Menu Card's interventions, assumes, in line with Cohen's index, the following classification: <0.20 = no or a negligible effect; 0.20 – 0.50 small effect; 0.50 – 0.80 = medium effect; >0.80 = large to very large effect.

This disagreement about how exactly an ES should be interpreted does not make matters any easier. To make matters worse, there are several other complications. It has been shown that ESs are not independent indices. Cheung and Slavin (2016, p. 283; see also Slavin & Smith, 2009), for instance, have demonstrated that the strength of ESs depends on a number of study features. They "are roughly twice as large for published articles, small-scale trials, and experimenter-made measures, than for unpublished documents, large-scale studies, and independent measures, respectively. In addition, effect sizes are significantly higher in quasi-experiments than in randomized experiments." Others (e.g., Gorard & Chen, 2025; Jones, 2019a, 2019b; Simpson, 2017; Wrigley, 2016), point to even more confounding influences, such as age/grade/phase and subject. Despite these unsolved difficulties, the Ministry decided to focus on ES in terms of months' learning gains (which is also provided for each intervention by the EEF). The effect size then indicates how many months of learning gain were achieved on average by students to whom the intervention was applied, compared to similar students when this intervention was not used. The Ministry warns users, however, that the toolkits don't make definitive claims about what works to improve learning outcomes at a particular school; they do offer high-quality information on what is likely to have a positive effect based on existing evidence. They don't provide ready-made answers and should therefore not be used as a stand-alone tool. The teachers should always take their expertise, judgment and context (including their specific students) into account, as well as the intervention's costs (NKO, 2026). (As already mentioned, according to the EEF (2023), selections should be based on a combination of its average impact on attainment; the strength of the evidence supporting it; and its cost – which will undoubtedly be a very difficult, if not impossible, task for most school leaders and teachers.) According to Katsipataki and Higgins (2016), the toolkits provide information not about "what works" but about "what has worked"; they show what has, on average, been

(or not been) effective, across the studies included. In order to make findings more accessible and understandable to school teachers and educational practitioners, the EEF has converted the ESs to months of additional gains in attainment. An ES of -0.04 – 0.04 equals 0 months' progress; 0.05 – 0.09 = 1 month; 0.10 – 0.18 = 2 months; 0.19 – 0.26 = 3 months; 0.27 – 0.35 = 4 months; 0.36 – 0.44 = 5 months; 0.45 – 0.52 = 6 months; 0.53 – 0.61 = 7 months; 0.62 – 0.69 = 8 months; 0.70 – 0.78 = 9 months; 0.79 – 0.87 = 10 months; 0.88 – 0.95 = 11 months; 0.96 – 1.00 = 12 months (EEF, 2023; MinOCW, 2025a).

The interventions presented in Table 1 are the ones selected by the Ministry in 2021. The data in columns 2 and 3 are based on an analysis of both toolkits. Column 2 shows the range (minimum – maximum value) of ESs for each intervention. Column 3 shows the mean effect of an intervention in terms of an ES (Cohen's d) and (in parentheses) the number of month's progress. As of 2025, the EEF is in the process of updating the toolkits, based on an improved analysis method. New studies were added, studies published before 1990 were removed, as were studies with less than 30 students. The total number of studies included in the toolkits amounted to nearly 3,000. As a result, there are – according to the EEF – some minor changes in the estimates of effect and strength of evidence for interventions. The composition of the Menu Card remained unchanged, though. In columns 4 and 5 of Table 1, the range and means of the (partly) updated toolkits are shown. To give a point of reference, column 6 shows the mean effect found by Hattie (2009) on the basis of 800+ meta-analyses.

Table 1. Effect sizes (Cohen's *d*) for evidence-based interventions (total) and additional months' progress ( )

	NEP Menu Card/ EEF Toolkits 2021		NEP Menu Card/ EEF Toolkits 2025		Hattie (2009) [e]
Intervention	**Min. – max.**	**Mean**	**Min. – max.**	**Mean**	**Mean**
Pre- and early-school interventions [a]	0.15 – 0.55	0.38 (+5)			0.45 (+ 6)
Extending school time [b]	-0.14 – 0.40	0.11 (+ 2)	-0.77 – 1.33	0.19 (+ 3)	
Summer schools [b]	0.00 – 0.43	0.18 (+ 2)	-0.57 – 1.83	0.25 (+ 3)	0.23 (+ 3)
One-to-one tuition	-0.06 – 0.70	0.37 (+ 5)	-0.93 – 3.47	0.40 (+ 5)	
Individualized instruction	-0.07 – 0.41	0.19 (+ 3)	-1.55 – 8.45	0.27 (+ 4)	0.23 (+ 3)
Small group tuition	-0.08 – 1.61	0.31 (+ 4)	-1.06 – 1.78	0.28 (+ 4)	0.49 (+ 6)
Direct instruction [c]	0.51 – 0.83	0.33 (+ 4)			0.59 (+ 7)
Peer tutoring [b]	-0.06 – 1.05	0.37 (+ 5)	-0.62 – 4.22	0.50 (+ 6)	0.55 (+ 7)
Feedback	0.20 – 0.97	0.63 (+ 8)	-1.57 – 2.75	0.48 (+ 6)	0.73 (+ 9)
Mastery learning	0.04 – 1.64	0.40 (+ 5)	-0.73 – 2.60	0.45 (+ 6)	0.58 (+ 7)
Reading comprehension strategies [b]	0.10 – 0.74	0.45 (+ 6)	-0.83 – 5.92	0.54 (+ 7)	0.58 (+ 7)
Spoken language interventions [b]	-0.14 – 0.91	0.37 (+ 5)	-0.86 – 4.54	0.45 (+ 6)	
Well-being interventions [d]					
Sports activities [b]	0.10 – 0.80	0.17 (+ 2)	-1.39 – 3.36	0.16 (+ 2)	
Arts participation	0.03 – 0.77	0.15 (+ 2)	-1.84 – 1.82	0.25 (+ 3)	0.35 (+ 4)
Metacognition and self-regulation [b]	-0.02 – 0.90	0.54 (+ 7)	-0.42 – 5.80	0.69 (+ 8)	0.69 (+ 8)
Collaborative learning interventions	0.13 – 0.91	0.38 (+ 5)	-2.11 – 3.51	0.45 (+ 6)	0.41 (+ 5)
Class size reduction [b]	0.12 – 0.34	0.19 (+ 3)	-0.26 – 1.79	0.08 (+ 1)	0.21 (+ 3)
Teaching assistant interventions	-0.15 – 1.50	0.08 (+ 1)	-0.26 – 1.62	0.35 (+ 4)	
Facilities and conditions [d]					
Parental engagement [b]	-0.14 – 0.65	0.22 (+ 3)	-2.23 – 1.92	0.28 (+ 4)	0.51 (+ 6)
Digital technology [c]	-0.15 – 1.13	0.29 (+ 4)			

Notes. Sources: Driessen (2021d); EEF (https://educationendowmentfoundation.org.uk/); Hattie (2009). [a] In the 2025 update this intervention was divided into 10 (sub)interventions. [b] This intervention was updated in 2025: new studies were added, studies published before 1990 were removed, as were studies with less than 30 students (EEF, n.d.-c). [c] Not in the EEF Toolkit anymore. d Not in the EEF Toolkit. e Hattie (2009).

Before discussing the table's contents, it is relevant to say something about the sources on which they are based (Driessen, 2021d). The 2021 data mainly come from meta-analyses, and, to a much lesser extent, single studies, while the updated 2025 data come from single studies. All this makes comparing the total number of the underlying studies nearly impossible. In addition, numerous publications date from before 2000, and thus are more than 25 years old. The oldest publication is even from 1948, that is, more than 75 years ago. Moreover, most studies themselves were conducted years earlier than their results were published. Furthermore, several publications are mentioned more than once and thus count double. The question, therefore, is how relevant those findings still are today. To give an impression, the year range of the references for Feedback was from 1961 to 2018. Only 36 percent dated from after 2000. In 2025, the EEF launched an updated version of the toolkits, starting with 10 interventions (EEF, 2025). While this certainly meant an improvement, the fact remains that many studies are still at least 35 years old; for Feedback this is even 27 percent.

What is striking is that the range of the ESs has dramatically increased between 2021 and 2025. In 2021, the ESs nearly all fell between -2.00 and +2.00, which is regarded as normal. In 2025, however, there were a few ESs greater than -2.00 in the table, while almost half of the ESs were greater than +2.00. In fact, one may wonder what the reliability of an ES of, for instance, 5.92 and even 8.45 is. Such extreme values are often indicative of poor study quality. It's most likely justified to interpret such implausible values as outliers, that should be excluded from the analyses. In addition to the range, the mean (total) ES of nearly all interventions has also increased. Why both changes have occurred is not clear. There does not seem to be a relation with the stricter selection criteria in 2025. If the EEF results are compared with those of Hattie (last column, Table 2), there appear to be several differences, though they are mostly not very large. It's also notable that two interventions from the Menu Card are not included in the EEF Toolkit: Well-being interventions and Facilities and conditions. This means that Menu Card users don't have comparable information available for the selection of these two interventions as for the other interventions.

A big problem of meta-analyses is that they lump all studies together. According to, for instance, Biesta (2010), Hattie (2009) and Kraft (2020), the practical application of any educational intervention should take into consideration the student, teacher and context (setting). Though the EEF (2023, n.d.-a) acknowledges and even stresses this caveat, still, the studies collected for the toolkits hardly make any distinction. Therefore, when school leaders, teachers, policymakers and researchers are looking for an intervention that works, the "evidence-based" interventions presented in the toolkit will surely offer not much practical help. The danger is that teachers will only consider the strength of the summarizing effect sizes presented (Wrigley, 2016).

In Table 2, a breakdown of subject (reading, math) and school phase (primary, secondary) is made for the EEF Toolkits. Despite this improvement, this for each intervention still leaves an amalgam of studies differing in terms of, for instance, age and grade, target group (general or specific, like disadvantaged students, gender, special education students, gifted students), individual or group approach, content and instructional design, intensity and duration, budget, type of outcome (e.g., observation, teacher assessment, formal test), research methodology and quality (e.g., experiment, quasi-experiment), sample size, and (statistical) analysis techniques. The core of Table 2 concerns the average ESs. In its computation of the average ESs, the EEF departs from the rule that for each intervention there should be at least 10 underlying studies. If there are fewer, no mean ES is computed. Oddly enough, this has not been applied to breakdown cells regarding subject x school phase. Therefore, the number of references per cell is also shown. When compiling this table, I discovered that many ESs on the EEF website were not correct. They are therefore not presented in Table 2 and are indicated with a b. (Early January 2026, in a professional blog, I commented on the EEF's ES of Feedback. For this, I used the values available from the EEF website. In all four cells of subject x phase, the average ES was 0 – which I found rather strange. Therefore, I checked a number of other interventions and found that there were at least 10 interventions where something seemed to be not quite right. On January 8, I informed the EEF, who, on January 12, replied that they had not been aware of the errors and would correct them. Which, on April 28, i.e., more than three months later, still wasn't done. Remember that the EEF website is an important resource, especially for school leaders and teachers (which also includes Dutch users). To give an impression: in 2024, the EEF resources were downloaded 1.3 million times, that is, more than 3,500 per day (EEF, n.d.-b). When asked, the EEF replied that school leaders and teachers do not read such "technical details", only researchers do. Is this really so? Every school leader should want to know whether the presented (total) ES also applies to their school level and subject. If a school leader were not interested in this, it would be a disqualification for their profession. Apart from all this, it may seem remarkable that no one noticed these important errors before, and that it took the EEF so long to correct them (which they actually didn't), nor didn't even bother to put a warning on their website. All this begs the question, therefore, how trustworthy the other coefficients are.)

The total number of references may seem large, but when a break-down is made according to subject x school phase, the resulting number is often (extremely) small, sometimes only 1 or 3 (e.g., Small group instruction; Mastery learning; Class size reduction). This means that selecting such "evidence-based" interventions is in no way justified. Moreover, for several interventions there are differences in ES according to subject and school phase. This means that the choice for such an intervention is only substantiated for that specific combination of subject and phase where the ES

is considered sufficiently strong. All in all, the data in both tables suggest that the choice of many interventions based on the presented ESs, the number of underlying studies and the month's progress does not seem justified at all.

Table 2. Effect sizes (Cohen's d) and number of underlying publications ( ) for evidence-based interventions (total, and broken down by subject x school phase)

Intervention	Mean	EEF 2025			
		Reading		Math	
		Primary	Secondary	Primary	Secondary
Pre- and early-school interventions [a]					
Extending school time	0.19	0.26 (40)	0.17 (13)	0.29 (13)	0.18 (10)
Summer schools [b]	0.25				
One-to-one tuition [b]	0.40	(92)	(6)	(12)	(3)
Individualized instruction [b]	0.27	(22)	(7)	(43)	(61)
Small group tuition [b]	0.28	(32)	(10)	(13)	0.05 (1)
Direct instruction [c]					
Peer tutoring	0.50	0.44 (52)	0.27 (12)	0.40 (31)	0.55 (20)
Feedback [b]	0.48	(30)	(25)	(12)	(10)
Mastery learning [b]	0.45	(8)	0.35 (1)	(15)	(19)
Reading comprehension strategies	0.54	0.54 (90)	0.52 (77)	-	-
Spoken language interventions	0.45	0.40 (59)	0.29 (19)	-0.02 (4)	-
Well-being interventions [d]					
Sports activities	0.16	0.14 (38)	0.45 (8)	0.16 (49)	0.20 (14)
Arts participation [b]	0.25				
Metacognition and self-regulation	0.69	0.50 (69)	0.67 (40)	0.69 (48)	0.71 (61)
Collaborative learning interventions	0.45	0.32 (40)	0.29 (16)	0.45 (35)	0.28 (27)
Class size reduction	0.08	0.08 (33)	0 (4)	0.12 (28)	0.45 (7)
Teaching assistant interventions [c]	0.35	(43)	(9)	(9)	(3)
Facilities and conditions [d]					
Parental engagement [b]	0.28				
Digital technology [c]					

Notes. Source: EEF (https://educationendowmentfoundation.org.uk/). a In the 2025 update this intervention was divided into 10 (sub)interventions. b No (correct) ESs available in the EEF Toolkits. c Not in the EEF Toolkit anymore. d Not in the EEF Toolkit.

3.3 More criticism

The previous section touched upon some criticisms of the EEF Toolkits and NEP Menu Card, but there are many more. Some of these are addressed below.

3.3.1 Meta-analyses

Some criticism of the use of meta-analyses and meta-meta-analyses has already been mentioned before. Wrigley (2016) lists a few more points, in particular regarding Hattie's (2009) meta-meta-analysis (the "Bible" of evidence-based education): the studies synthesized are mostly from the USA; they are often 30–50 years old; many are limited by narrow outcome measures which do not reflect important educational aims; most are based on early literacy or numeracy. In addition, he points to other aspects which make the computation of average ESs problematic: there is a lot of variety in duration of each intervention; a diversity of outcomes is lumped together, such as reading, science, behavior; average ESs decrease dramatically with age; no difference is made between very specific and very broad groupings; interventions focus on very broad issues, such as "home" and "personality"; sometimes an ES is interpreted "as compared to a control group" and at other times "as compared to the same students before the study started".

Coe (2002) warns that in comparing (or combining)

effect sizes from meta-analyses, it is necessary: to consider whether they relate to the same outcomes; whether those outcome measures are derived from similar instruments and populations; to compare only like with like in terms of the treatments used. "…the same name is often given to interventions that are actually very different… It could also be that ... the actual implementation differed, or that the same treatment may have had different levels of intensity in different studies. In any of these cases, it makes no sense to average out their effects." (Coe, 2002, p. 10).

According to Gorard and Chen (2025), there is another, perhaps even bigger problem, namely that meta-analyses are conducted without consideration of the quality or trustworthiness of the underlying evidence. Many studies are quasi-experimental or correlational, only a few are true experiments (RCTs), but there is no guarantee that the studies in question are of sufficient, let alone high, quality (cf. Driessen, 2025; Gilleece & Clerkin, 2025; Styles & Torgerson, 2018). As a consequence, the ES of a poor quality study is treated as equivalent to that of a very good study. "This does not make any kind of sense." "…effect sizes cannot reasonably be aggregated or averaged, except where all of the studies involved were of the same design and quality" (Gorard & Chen, 2025, p. 4116).

3.3.2 Determining the target group

Schools with disadvantaged students did receive additional funding under the NEP, in addition to the supplementary funding all schools already received to combat the disadvantages caused by COVID-19. Moreover, these schools also receive substantial additional budgets under the Education Disadvantage Policy (Driessen, 2022). These schools were therefore receiving additional funding from (at least) three sources. All in all, these are enormous sums. A major concern is whether these billions of euros have been properly allocated. It has been demonstrated that the criteria used in allocating these funds are not truly valid (Driessen, 2017). This implies that many false-positive and false-negative target group children have been identified, respectively have not been identified; in other words, for many children, the schools have wrongly received additional funding, or indeed no funding at all.

3.3.3 Months of progress

The Ministry has chosen to focus on months of progress in order to identify what works best. Therefore, ESs were converted into months of progress by the EEF. Serious objections have long been raised to simply converting effects into months of (additional) growth; a procedure that is therefore strongly discouraged (Baird & Pane, 2019). Growth is not linear; there are differences by grade level/age, subject/discipline, and test/measurement instrument (Dadey & Briggs, 2012). This risks lumping everything together and comparing apples and oranges (Simpson, 2018). Moreover, growth arises not only from schooling, but also from maturation and extracurricular influences (Luyten et al., 2017). And growth also depends on the starting level, the degree of initial disadvantage. And, of course, such a conversion is insignificant for non-cognitive measures, such as well-being or motivation (Jones, 2019a, 2019b). There is yet another problem, namely the reference group. Didau (2017) gives the following example. The Feedback intervention results in 8 months of extra progress. But the question is: compared with what? Not giving feedback? But every teacher gives feedback! However, all this raises a much larger issue: what additional progress are we actually talking about? The Menu Card includes the months of additional progress from the EEF Toolkits, but does it apply to language, arithmetic, motivation, behavior, or health? And does this progress apply to all domains, or is there even progress in some areas and regression in others? Such a specification is, of course, essential when school leaders and teachers need to make a responsible choice among interventions aimed at the specific disadvantages of their students. Since these disadvantages manifest themselves in various areas, multiple interventions will undoubtedly be necessary.

3.3.4 Practical issues

At the start of the NEP, the Netherlands Court of Audit (Algemene Rekenkamer, 2021, 2022) listed several more practical criticisms:

- The objectives of the NEP are not sufficiently clarified, making it impossible to assess after the program's completion whether this huge investment has the desired effect. Billions of public money are being spent on a program whose achievement of its objectives can never be assessed.
- The Minister has established a national monitoring and evaluation program. This includes an implementation monitor that provides insight into the use of interventions by schools. The results monitor provides insight into the extent and distribution of student delays caused by the coronavirus crisis and how these delays develop over time. Finally, the Minister wants to use an impact assessment to provide insight into the effectiveness of 5 to 10 "promising" interventions. However, according to the Court of Audit, it is highly questionable whether and how this information will provide insight into the results of the NEP. This is because, within the monitor no link can be made between resources, types of interventions, and the results achieved at the school level.
- Because the Minister and the Parliament considered it urgent to resolve the COVID-19 delays, NEP funds were added to the existing lump-sum funding. As a consequence, the Minister could not, within the legal framework, impose specific conditions on what and when schools had to spend this money. He also could not reclaim the money if schools make different choices or do not spend it within the two-year term.

However, when informing schools and Parliament, the Minister gave the (therefore: false) impression that schools are obligated to spend the money within the term of the NEP, and solely on interventions on the Menu Card (cf. Brands & Schijns, 2024).

- Many schools want to spend their money on additional staff. This creates a high demand for teachers. Due to the already existing teacher shortage, these are scarce, however. In a survey, 41 percent of the school leaders indicated that they had not been able to fill all vacancies created by the NEP. Moreover, schools with a greater risk of disadvantage had more difficulty attracting and retaining teachers than other schools. As a result, schools are effectively competing with each other for staff, which is detrimental to schools with the greatest risk of disadvantage.
- School leaders indicate that they do not use the Menu Card as a tool for selecting interventions. Instead, they view the Menu primarily as an accountability tool. School boards have to indicate in their annual reports on how much funding their schools have allocated for each intervention. As a result, school leaders did not arrive at different choices based on the Menu, but instead attempted to justify the choices already made as interventions within the Menu. Schools were inclined to continue on their current path and invest more in it, rather than making new choices.
- The Menu Card contains too many interventions. Furthermore, the interventions are formulated so broadly that the Minister offers little help to schools. They could interpret virtually any existing plan as an intervention from the Menu.
- It is not easy to trace the budgeted funds in the school plans. Striking is that schools primarily want to spend money on interventions that, according to the Menu, were relatively expensive and less effective than other measures listed, such as hiring teaching assistants and reducing class sizes.

3.3.5 The Toolkits' studies

The EEF (2023) acknowledges that there are several hooks and eyes regarding the selection of the studies used to build the toolkits (remember, however, that in 2025 a number of interventions have been updated):

- Only a few studies have taken place recently. This might mean that the research is not representative of current practice.
- Many studies included are not RCTs; there is a risk that results are influenced by unknown factors that are not part of the intervention.
- Numerous studies took place in conditions distinct from regular school practice.
- While all studies included take place in real classrooms (as opposed to "lab studies"), studies that are delivered by researchers rather than teachers tell us less about how an intervention will be implemented.
- Many studies were not independently evaluated. Evaluations conducted by organizations connected with the intervention, for example, commercial providers, typically have larger impacts, which may influence the overall impact of the intervention.
- There is a large amount of unexplained variation between the results included in the topic. All reviews contain some variation in results, which is why it is important to look behind the average. Unexplained variation (or heterogeneity) reduces the certainty in the results in ways that the EEF has been unable to test by looking at how context, methodology or intervention is influencing impact.

In the initial phase of the NEP, Driessen (2021a, 2021b, 2021c) presented a number of other comments:

- The interventions encompass diverse (practical) implementations; they concern a very broad, abstract concept; it is not always clear what exactly is meant.
- There is little or no targeted research available, that is, research (and evidence) from the COVID-19 context, focusing on school closures, homeschooling, distance learning, and not only cognitive, but also psychosocial, emotional and behavioral outcomes. The research available has a different objective and target group: interventions are generally designed to prevent, reduce, or eliminate disadvantages for children from lower incomes and ethnic backgrounds.
- The research used is primarily from the USA and, less so, the UK; Dutch research is almost completely lacking. The context (and school system) therefore may be absolutely incomparable.
- The (meta-)analyses presented by the EEF show considerable overlap, especially when conducted by the same author.
- Research is often methodologically flawed.
- There appears to be little consistency: there are positive, null, and negative effects, from strongly positive to strongly negative.
- Any positive effects in a study appear to be overreported; little to no attention is paid to negative and null effects in the same study. Any effects apply only to specific target groups (e.g., socioeconomic environment, ethnicity, age/grade/phase) and subject/content domains (e.g., language, math, self-confidence, motivation).
- More recent studies often show weaker effects than older studies.

According to Wrigley (2016, 2018), school heads and teachers are likely to focus on the months of added progress figures on the EEF website, with the imminent danger that this will lead them to jump to conclusions. There will not be many teachers (and researchers, for that matter) who are familiar with the underlying methodology and aware of its many complications (cf. Simpson, 2017; Gorard & Chen, 2025), such as:

- the computation of averages on the basis of dissimilar studies;
- the selectivity of sources;
- the great variety of individual scores that underlie the average, many of them qualifying or even contradicting the average;
- the context-dependency of many interventions, often rooted in a specific usage;
- the ambiguity of the precise nature of the interventions;
- the exclusive focus on attainment, with no consideration of other educational goals;
- the lack of attention to the needs of the students the toolkit is supposed to help, i.e., the socioeconomically disadvantaged students.

The RED-team Education (RED-team Onderwijs, 2021; a group of critical education experts) has the following objections to the NEP:

- Quantifying and identifying learning gaps that have arisen during the COVID-19 pandemic is extremely difficult, if not impossible.
- At most schools, especially in secondary education, no one knows which students are lagging behind, what they are lagging behind in, how extensive these gaps are, and what the best intervention is to eliminate them. Therefore, the amount of money allocated is not linked to a sound analysis of the problem. This carries a significant risk of wasting money, time, and energy.
- To prevent this risk, schools must develop a plan based on their own inventory and a proven intervention to eliminate learning gaps. This bureaucratic procedure suggests control, but is barely feasible and, above all, costs schools bureaucratic energy. Moreover, there exists no "proven approach" for the present complex problem.
- The real limitation lies in the quantity and quality of teachers and their lack of time to do their work properly. This shortage of teachers, even without the pandemic, resulted in serious learning gaps for students. This is precisely what schools cannot address with this incidental funding. Schools will be forced to purchase their services from employment agencies and homework institutes. It is highly questionable whether this will yield the intended boost in quality.
- €5.8 billion is a lot of money. Previous experience with making large sums available without concrete targets shows that this results in reputational damage, followed by a predictable political reflex. In the next round of budget cuts, which is guaranteed to come, education will be presented with the bill. This yo-yoing with financial resources—first generously donating, then cutting back—is hindering the resolution of the structural quality problems in education.

RED-team member Kirschner (Scheerens & Kirschner, 2021) calls the Menu Card a "shotgun blast." The interventions are so broad and general that they're essentially worthless. The Menu doesn't include ready-made meals or detailed recipes. As a result, anything is possible. He summarizes the policy as a "panic reaction" at the Ministry. "They thought: There's a crisis, so there must be backlogs. We're just throwing a bag of money at it, without first properly assessing what's going on. And then hoping it will help." (Brands & Schijns, 2024, pp. 7, 12).

Van den Broek et al. (2025) conducted a large-scale survey among school leaders and teachers in primary and secondary education, with which they tries to map out how the entire NEP process had progressed.

Overall, the **preparation** of the NEP was qualified as positive. Schools and partners appreciated the clear formats, the freedom of choice, and the extra resources. The selected interventions were, in most cases, well-suited to the identified educational needs. In primary education, between 5 and 8 interventions were selected. Schools repeatedly cited time pressure, workload, and staff shortages as obstacles. Table 3 provides an overview of the most frequently selected interventions. In light of the teacher shortages, it is striking that the two most frequently selected interventions involved deploying additional staff and support. Also, interventions focusing on the socio-emotional and physical development of students were relatively frequently selected. Regarding potential effects, it is notable that the results of two interventions are unknown. In general, the number of months of predicted extra gains, with a range of -2 to 8 weeks, is not particularly high. This raises the question to what extent this criterion played a role in the choice of an intervention, or whether other factors were decisive.

Across the board, schools consider the **implementation** of interventions within the NEP framework to be successful. Most activities were carried out as planned, and collaboration within teams and with partners was perceived as strong. At the same time, staff shortages, workload, sick leave, and temporary funding were common obstacles to continuity and sustainability. To address staff shortages, extensive use was made of external parties, representing 19 percent of the budget in primary education and 17 percent in secondary education. These parties were primarily deployed for summer and spring schools, cultural education, and sports activities; in secondary education, they were also deployed for professional development, guidance, or organizational support. In primary education, success factors included team collaboration, motivation, and effective alignment of interventions with student needs; in secondary education, these included a strong and learning team, flexibility, and student engagement.

The **results** were generally also assessed as positive. According to the school leaders, the NEP goals (recovering from learning delays and strengthening social-emotional well-being) have been largely or fully achieved in many schools. At the same time, mental health, vulnerable home situations, and motivation remain important concerns. In primary education, it helps when interventions are closely aligned with the actual learning delays and when schools

operate evidence-informed. A comprehensive approach (7–10 interventions) is also associated with greater success, as is increased external hiring (up to approximately 30%). In secondary education, concerns remain regarding vulnerable students, mental health, and basic skills. The highest results are reported for teaching assistants, class size reduction, well-being interventions, and sports activities. Team quality also appears to be a key success factor, while teacher shortages and limited implementation capacity hinder the achievement and maintenance of results. Concerns remain greatest in schools with many disadvantaged students and in urban areas. The impression exists that schools with many disadvantaged students have benefited no more or less from the NEP funds than schools with few such students. Regarding the results, it is important to note that these are solely assessments obtained from school leaders via questionnaires, and not objective test results. Another point of concern is that a very large part of the budget has been spent on the (expensive and temporary) use of external, commercial parties.

Reflections on the NEP by the school leaders reveal a strong sense of appreciation across all sectors for what has been initiated with the NEP funds: recovery, innovation, and increased collaboration. At the same time, the common thread is that temporary nature, workload, and bureaucracy have hampered structural integration. The key lessons revolve around the value of sustainable investments in people, teams, and collaboration, and the importance of well-being as a foundation for learning.

Table 3. Most chosen interventions according to school leaders

Intervention	Months' progress [a]	Primary school	Secondary school
Small group tuition	+4	85%	79%
Teaching assistant interventions	+4	79%	76%
Well-being interventions	?	70%	91%
Direct instruction	?	48%	0
One-to-one tuition	+5	47%	65%
Sports activities	+2	0	75%

Notes. Source: Van den Broek et al. (2025). a Range: -2 – 8.

Haelermans and Havermans (2026) performed a number of quantitative analyses based on regional, national and international datasets (also see Van Vugt et al., 2025a, 2025b, 2025c). Data concerned both cognitive (reading, spelling, and math test scores) and non-cognitive outcome measures (creativity, curiosity, critical thinking, and problem-solving abilities). The problem here is that most analyses focused on developments over time (pre-COVID-19, COVID-19, and post-COVID-19), while there were no data on the NEP, i.e., interventions, available. Therefore, it is not possible to study relations between the two, let alone the causal impact. Though there was quite some variation between school years, sectors, grades and subjects, Haelermans and Havermans conclude that students are slowly catching up after the COVID-19 pandemic, although this applies to a greater extent to primary education than secondary education. In addition, there are gender and social milieu differences: boys do better than girls, and students with high-educated parents worse than students with middle- or low-educated parents. According to Haelermans and Havermans, this might be explained by the fact that many policy interventions primarily target disadvantaged students, although it is questionable whether this is actually the case). They also argue that parents with higher education levels work, on average, more hours face greater difficulties in combining working from home, taking care of the children, and offering homeschooling, which could negatively affect their children's school performances. In secondary education, the situation differs somewhat from primary education. Learning losses persist after the pandemic compared to during the pandemic. Regarding cognitive skills, they conclude that they have also declined during and after the pandemic compared to pre-pandemic; this holds especially for primary education. For secondary education, the situation is more complex: some non-cognitive skills are higher post-pandemic than during and before COVID-19, while others are lower. Impact analyses of policy interventions show some promising results for student performance, although not all interventions have reached the desired effects. "In sum, we see the resilience of students in bouncing back after the pandemic, but in many respects, the Netherlands are still not at the same levels as before the COVID-19 pandemic." (Haelermans & Havermans, 2026: 126).

Based on the above, the Deputy Minister of Education (Tweede Kamer, 2025) concludes in his letter to the House of Representatives that changes in students' school careers during and after the pandemic have remained limited in primary education. The impact on school careers in secondary education has been greater than in primary education:

- The learning delays resulting from the COVID-19 pandemic have been recovered for students in primary

education. Learning performance has now returned to pre-pandemic levels.
- Learning performance in secondary education shows a mixed picture and requires continued attention.
- The decline in equal opportunities in primary education due to the pandemic has been reversed.
- Learning is not separate from well-being. There is visible improvement in student well-being, and school leaders and teachers now indicate that they are paying more attention to well-being.

According to the Deputy Minister, there is a connection with the design of the NEP: a quick start of the program with concrete measures; targeted support for the most vulnerable students; careful monitoring; explicit attention to well-being; and space for schools to choose from evidence-informed interventions within the Menu.

Oddly enough, the State Secretary's letter is completely devoid of any self-criticism. Everything is formulated positively. It doesn't address the concerns raised at the outset of the NEP, nor the consequences thereof. It also doesn't address what the Ministry could perhaps have done better in selecting truly evidence-informed interventions.

5. An international perspective

This study provides an in-depth investigation into how the Netherlands managed the impact of the COVID-19 pandemic on education. The focus is on the reliability and validity of the approach chosen by the Ministry of Education, particularly regarding mandatory evidence-based interventions. It is now relevant to examine how other countries have addressed this issue. A cross-country comparative study was recently published, examining how other nations respond to the negative effects of the pandemic (Volente et al., 2026). A brief summary follows, focusing on several Western European countries: Belgium, Germany, England, France, Italy, the Netherlands, and Sweden. In addition to the aforementioned publication, references were also made to, among others, Schnepf et al. (2024).

This overview study shows that there are many similarities and differences among the countries considered. For example, the degree of centralization and autonomy is a relevant factor in choosing a specific approach. Countries such as England and the Netherlands, have semi-decentralized education systems: much is determined by the central government, but schools (specifically, school boards) retain considerable freedom in implementation. France has a highly centralized system, with the Ministry of Education exerting significant influence over policy implementation. This results in a standardized approach, aimed at uniform quality and equal opportunities for all students. Germany has delegated education policy to its 16 states. While there are similarities, each state has its own policies and powers. Centralization has advantages, but also disadvantages. The advantage is that decisions can be made and action taken more quickly in a crisis. The disadvantage is that a uniform approach is often chosen, which offers fewer options for customization, for example, with regard to regional and sectoral differences, and coordination with the specific needs of various target groups.

The choice of a particular approach to the COVID-19 pandemic was initially a political one, usually advised by various educational and healthcare institutions. One problem was that the pandemic spread rapidly, and the government had absolutely no experience with it. The government was, so to speak, at its wit's end. In the vast majority of countries, schools were closed to prevent further infections. However, there were significant differences in the duration of these closures. In France, it lasted only 8 weeks, while in Italy, it lasted almost double that, namely 15 weeks. Germany fell somewhere in the middle, with 8 to 11 weeks, as did England, with 12 weeks. In the Netherlands, the closure duration varied by sector; in primary education, it was 11.5 weeks and in secondary education, 16 weeks. The Swedish government's decision to allow primary education to continue as normal was completely different.

There are significant differences in the budgets made available to combat the COVID-19 backlog. In the Netherlands, this amounted to an unprecedented total budget of €8.5 billion, of which €5.8 billion was for primary and secondary education (nearly 2.5 million students). The available budgets in other countries were considerably smaller. England, for example, spent only £4.3 billion, of which £1.3 billion (or 30%) was due to underspending or existing budgets, resulting in a total budget allocation of only £3.0 billion (Farquharson et al., 2021). Germany allocated €2 billion to education (€93 per student). De Witte and Smet (2021) present an overview of what was spent per student in several countries: Belgium (Flanders) €486; Belgium (Wallonia) €44; Italy €261; The Netherlands, by far, takes the cake with a whopping €2,795.

Due to the speed at which the pandemic developed, governments in every country were completely surprised by what was coming their way. This situation was not only unexpected and rapid, but also completely new, and this applied to all countries. Specific policies therefore had to be developed from the ground up and at breakneck speed. A sense of urgency was felt in every country; there was no time for thorough preparation – according to the relevant ministries. Consequently, the chosen approaches were not really based on theoretical foundations; the solutions were therefore mainly ad hoc, the implementation of which often depended on the education system, previous education policy, and available knowledge and tools. Several countries did, however, seek to integrate evidence-based interventions. What follows is a brief overview per country.

Belgium (Gambi & De Witte, 2026). Belgium has a highly decentralized governance system, with powers at the federal, regional, and language levels; in addition, there is the principle of freedom of education. Initially, there was considerable collaboration between these levels, but gradually regions and communities opted for their own

measures. These focused on three areas: digitalization, mental health support, and learning recovery. Ultimately, various interventions were implemented in all three language communities, notably summer schools; an action plan to combat school dropout (career guidance and job-related training in vocational and technical secondary education); a monitoring system focusing on early school-leavers' outcomes, socio-economic backgrounds, and academic progress; hiring and training teachers and providing counseling and support services for students. The extent to which all these measures have been successful remains unclear due to a lack of suitable data.

England (Badunenke, 2026). Various measures were implemented in England. The National Tutoring Program focused on one-to-one and small-group tutoring for disadvantaged pupils. Recovery Premium Funding offered targeted academic support to schools, with additional resources for the most affected students. The Education Recovery Plan allocated funding for teacher training and professional development to improve the quality of instruction. Summer schools were established for secondary students, focusing on core subjects and well-being. Investments were also made in the Oak National Academy, which offered online lessons to supplement classroom instruction. Schools were advised to implement flexible learning strategies, including extended school hours and catch-up sessions.

France (Maurya & Moulin, 2026). The centralized education system enabled the French government to quickly implement a uniform national policy. However, a challenge was accommodating regional differences and diverse target groups. Various programs were implemented in different phases and at different levels. During the pandemic, the focus was on emergency measures, such as deploying digital learning platforms and keeping schools open as much as possible. Structured learning activities and tutoring sessions were also offered during school holidays to help students catch up on missed education. After the pandemic, the focus shifted to long-term recovery through digitalization in education.

Germany (Schult et al., 2026). Due to the high degree of autonomy of the 16 German states, it presents a unique case. Each state had its own regulations, which sometimes conflicted with those of other states. In general, most teachers and students quickly adapted to the new situation, primarily following a pragmatic approach (use what works). There was no uniform crisis management system for the 16 states. The policy and its practical implementation were a combination of state-specific and individual teacher-based actions. The scope and quality of remote instruction depended on the efforts of teachers, students, and their families, which varied considerably, especially for disadvantaged students. The focus was primarily on basic literacy and numeracy skills, both in the classroom and during school holidays. The most affected student groups received extra attention; efforts were also made to train and hire additional teachers. Besides cognitive skills, there was also a focus on strengthening social learning through schools and youth work. Due to the sovereignty of the federal states, policies regarding, for example, external tutoring, holiday courses, or social-emotional skills were implemented in very different ways and with different priorities. Responsibility for these was delegated to schools and teachers. Due to the perceived urgency, the measures taken could not be empirically validated, and cognitive and non-cognitive developments were not mapped in a uniform manner. This made it impossible to carry out an adequate evaluation and to assess the effectiveness of the measures.

Italy (Giancola & Salmieri, 2026). Italy's hybrid educational governance model situates it between highly centralized and decentralized systems, with the Ministry of Education holding the most authority and individual schools enjoying moderate pedagogical autonomy. The COVID-19 policy was implemented in two phases: preventive measures during school closures and, after the pandemic, remediation efforts. During the first phase, digital equity was central, with hardware provided to disadvantaged students and teachers trained in remote instruction tools, building on existing plans. Although schools received additional funding, there was no remedial programming to address disadvantages in 2020. Interventions implemented after 2021, such as selective in-person tutoring in underserved areas, depended on local needs but were not implemented on a larger scale, and impact research was absent. Consequently, the emphasis in Italy was on an individual-centered approach, ignoring acute, cohort-specific vulnerabilities. After the pandemic was managed, policy focused on modernizing educational infrastructure rather than addressing the accumulated learning gaps. Particularly lacking was attention to the social-emotional aspects of learning loss. Implementation of policy measures varied greatly by region and actually exacerbated existing inequalities. Everything was fragmented; local measures lacked the coordination to effectively address the problem at the national level. Interventions were based on existing, small-scale, and localized strategies. A comprehensive national program was lacking, as were impact evaluations.

Sweden (Wikström & Wikström, 2026). The Swedish education system is highly centralized, yet also shows some decentralization , particularly regarding budget allocation and educational practices. Sweden chose to keep schools open for children up to 15 years old, assuming that the risk of COVID-19 infection outweighed the potential negative social-psychological effects on young children. During this period, testing and monitoring of students continued, although absenteeism was common. Recent results from the international PISA study show that literacy and arithmetic performance declined sharply. Conversely, national research showed that overall performance remained stable during the pandemic. Overall, the picture is mixed. For older Swedish children, schools were closed and distance learning was implemented, requiring adjustments from both students and teachers. The government provided additional resources for this purpose. The performance of

these students does not appear to have suffered from the transition to distance learning. However, concerns remained about their mental health and social development. Research in this area also showed mixed results. In some cases, there appeared to be a correlation with students' socioeconomic background. The extent to which these trends are related to COVID-19 remains unclear.

What can we conclude from the above?

- All countries were completely surprised by the COVID-19 pandemic; they were totally unprepared for its consequences.
- School closures varied widely, from 8 to 16 weeks; in contrast to other countries, schools in Sweden remained open.
- Governments and Ministries of Education quickly took action to mitigate the expected negative effects of the school closures. Due to the perceived urgency, there was little time to develop an adequate plan. Often, work already underway continued.
- The speed of decision-making and implementation of the plans was closely related to the structure of the education system: in centralized countries, everything moved more quickly; in decentralized countries, there was more room for customized solutions.
- The size of the budgets made available varied enormously. For example, in Wallonia (Belgium), only €44 per student was allocated, compared to €2,795 in the Netherlands.
- The chosen interventions had many similarities. Often, the evidence-based paradigm was adopted — intentionally or not. The most common approaches were: summer schools; hiring and training teachers; one-on-one and small-group tutoring; distance learning; and digitalization.
- The emphasis was generally on basic literacy and numeracy skills. In some countries, but not all, attention was also paid to the social and emotional development of the students. Almost universally, children from socio-economically disadvantaged families received extra support.
- In only a few countries was there a long-term approach. Little attention was given to (future) evaluations and effect studies, resulting in insufficient or no monitoring of the programs, and no suitable data on the children's development being collected. Consequently, in almost no country could a relationship be established between the use of interventions and the results. In other words, in no country was there any effect study. Therefore, the effectiveness of all the implemented policies is completely unknown.
- In most countries, the cognitive and non-cognitive development of the students were assessed. This was done based on data from international studies that are essentially unsuitable for determining whether the observed developments are the result of policy. The broad conclusion is that there is some recovery in cognitive development, but students are not yet at pre-COVID-19 levels. Children from lower socioeconomic backgrounds are experiencing the greatest disadvantages. In non-cognitive areas (social and emotional development), significant problems remain and are expected to persist for some time.

6. Discussion

When the COVID-19 pandemic broke out in the Netherlands in 2020, the country was completely unprepared for its potential consequences. The Minister of Education hastily began developing a National Education Plan to address the anticipated backlog. This proved extremely challenging, as this was the first time such a disaster had occurred. First, he allocated an unprecedented amount of funding for schools to use. How he arrived at this amount remains unclear. He didn't know how to implement the plan in concrete terms; long-term school closures were an unknown phenomenon until then. The evidence-based approach, however, was intended to be the guiding principle. Because the Netherlands lacked a What Works database of scientifically proven interventions, he turned to the UK, where one had already been under development for some time. After translation and additions, over 20 evidence-based interventions were ultimately included in the Menu Card. Schools were required to choose from these and spend the funds within the plan's duration. Although the Minister knew that both conditions were legally untenable, he stuck with them—due to the sense of urgency he felt. Ultimately, however, he had to give in: schools were free to choose interventions and had to spend the extra funding in four, not two, years. He was convinced that with a huge sum of money and a list of proven interventions, everything would be fine. He assumed that he had made the right decision, both scientifically and practically, with the interventions adopted from the UK. What he forgot to do was check for any pitfalls in that choice. There were, unfortunately, numerous ones. Relevant questions such as "How evidence-based are these interventions, really?"; "Have all methodological problems been acknowledged and resolved?"; "Do the described outcomes apply to the current time, to the Netherlands, to the COVID-19 situation, and to the specific context and student population of a school?"; and "Can schools make a substantiated selection and apply it in their specific situation?" were not, or at least not sufficiently, addressed by the Minister. He was also in such a hurry that he neglected to set any concrete goals. Without such goals, it was impossible to determine whether the deployment of all that money had been effective. Experts summarily described this policy as a "panic reaction" and a "shotgun approach." From the very beginning, warnings were issued from all sides that, due to poor preparation, ill-considered choices and one-sided focus on "scientifically proven interventions", it would never be clear whether the policy had been successful,

which regrettably came true.

In this article, we have concluded multiple times that there is no evidence that the NEP has been effective. Although there was input, in the form of money and interventions, and output, in terms of the post-COVID-19 level of cognitive performance and socio-emotional characteristics, no causal link could be established between the two. The fact that this cannot be demonstrated does not mean that such a connection does not exist. Perhaps it does; we simply do not know.

In the Deputy Minister's final letter to Parliament (MinOCW, 2025), he seems to apologize for the lack of clarity regarding the effectiveness of the NEP. "The precise effect of the National Education Program is difficult to determine. After all, we cannot know what would have happened without the National Education Program" (Tweede Kamer, 2025, p.1). However, there are at least two reasons to assume that the NEP did not really contribute to the (partial) recovery. First, this article has made clear that the (one-sided) focus on evidence-based interventions, on effect sizes and months of additional learning gain, as it was applied here, cannot have truly contributed to that recovery. Confidence in it—again: as applied here—has proven unfounded, built on quicksand. Second, if there was any confidence in that evidence-based approach, despite this, it is undermined by the practice encountered. After all, schools have primarily opted for interventions with at most a weak impact, or interventions whose impact is unclear or unknown. They seem to have continued on their usual path, with a business-as-usual approach, and have seen the Menu Card as a way to justify their choices and the allocation of the allocated budgets afterward.

This article is about the Netherlands. The above also provides a glimpse into how other countries have dealt with the consequences of COVID-19. The Netherlands stands out positively in several respects, such as its exceptionally high budget; the national data collections and monitoring of developments; the choice of a range of diverse evidence-based interventions; and the extra attention paid to disadvantaged students. A highly relevant question (also for the future) is to what extent all of this has led to a greater impact and a better position for students in Dutch education than in other countries that have invested (considerably) less in COVID-19 policy. Unfortunately, this question will never be answered. The reasons for this are explained above.

The foregoing leads to the following recommendations to be better prepared for future calamities:

- Recently, November 2025, the evaluation of the NEP has been published. The results presented in that report only concern the short-term outcomes. To learn more about the long-term outcomes, it's necessary to follow the student cohorts and monitor how they further develop during their career in education and afterwards, in the labor market. In addition, the school leaders and teachers should also be followed, to check what they've learned from the NEP and how they used it afterward.
- Insofar applicable and useful, disseminate the evaluation's findings among schools, in such a way that schools can really benefit from it in their daily practice.
- Be much more critical regarding the use and even obligation of evidence-based information. It is evident that the information presented by the Ministry lacks knowledge of the subject and oversimplifies things.
- Create a database of truly proven interventions, applicable to the present, the Netherlands, and specific subjects and ages/grades. This requires a group of experts, from both academia and practice, who dare to examine "the evidence" independently and—above all—critically. Do not only look at the positive effects, but also at the null and negative effects. Compiling all sorts of studies on an intervention in a meta-analysis without specifying its conditions is pointless.
- Develop and thoroughly evaluate (new) interventions in high-quality research specifically tailored to a specific situation. There will undoubtedly be more pandemics in the future.
- Don't assume that what has been developed abroad (and in the past) is immediately and without further ado applicable to the Dutch context and school system.
- Take more time to properly prepare a policy and develop it into a more concrete plan. Clearly indicate what you want to achieve and how. A thorough analysis of the situation, therefore, is a sine qua non.
- Impose more requirements on schools regarding the allocation of the additional funds, the valid and reliable identification of the target groups and the choice of interventions.
- Better prepare and support schools in making and implementing their choices, for example, with a kind of flying squad.

From the very beginning, the NEP had significant caveats, which were communicated to the Minister, but he barely addressed them. Now, several years after the pandemic, students seem to be back on track to some extent, although the (e.g., psychosocial) consequences will continue to reverberate for a long time. The extent to which the current situation has anything to do with the NEP and the deployment of billions of euros in subsidies will always remain unclear. Therefore, we will never know whether the so-called evidence-based interventions were effective or not. A missed opportunity!

Author's contributions

The author conceptualized the study, performed the literature review and analyses, and drafted the manuscript.

Conflict of interest

The author declares no conflicts of interest.

Funding

The author received no funding.

Ethical statement

This study describes (secondary) data collected from schools, school leaders, teachers and, students in a range of countries. The author assumes that the original researchers obtained appropriate ethical approval and informed consent. All data in this article have been de-identified and only reported in a high degree of aggregation.

References

Algemene Rekenkamer. (2021). *House of Representatives' letter: Points for attention regarding the National Programme for Education*. https://www.rekenkamer.nl/documenten/2021/03/31/aandachtspunten-bij-het-nationaal-programma-onderwijs.

Algemene Rekenkamer. (2022). *Accountability audit of the Ministry of Education, Culture and Science: Report on the 2021 Annual Report*. https://www.rekenkamer.nl/documenten/2022/05/18/resultaten-verantwoordingsonderzoek-2021-ministerie-van-onderwijs-cultuur-en-wetenschap.

Badunenko, O. (2026). Academic resilience and policy reform in England. In L. Volante, K. De Witte, O. Giancola, L. Salmieri, & D. Klinger (Eds.), *Risk, resilience, and recovery across global education systems: The long-term impact of pandemic reforms* (pp. 15-29). Routledge. https://doi.org/10.4324/9781003608806-4.

Baird, M., & Pane, J. (2019). Translating standardized effects of education programs into more interpretable metrics. *Educational Researcher, 48*(4), 217–228. https://doi.org/10.3102/0013189X19848729.

Biesta, G. (2010). Why "what works" still won't work: From evidence-based education to value-based education. *Studies in Philosophy and Education, 29*(5), 491-503. https://doi.org/10.1007/s11217-010-9191-x.

Brands, N., & Schijns, V. (2024, January 29). Niemand weet waar de coronamiljarden voor het onderwijs zijn gebleven. *Follow the Money*. https://www.ftm.nl/artikelen/waar-is-het-npo-geld-gebleven.

Cheung, A., & Slavin, R. (2016). How methodological features affect effect sizes in education. *Educational Researcher, 45*(5), 283–292. https://doi.org/10.3102/0013189X16656615.

Coe, R. (2002, September 12-14). *It's the effect size, stupid: What effect size is and why it is important.* British Educational Research Association Annual Conference, University of Exeter, England. https://dradamvolungis.com/wp-content/uploads/2012/01/its-the-effect-size-stupid-what-effect-size-is-why-it-is-important-coe-2002.pdf.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates, Publishers. https://doi.org/10.4324/9780203771587.

Dadey, N., & Briggs, D. C. (2012). A meta-analysis of growth trends from vertically scaled assessments. *Practical Assessment, Research & Evaluation, 17*(14), 1–13. https://doi.org/10.7275/f2bm-6r59.

Dekker, I., & Meeter, M. (2022). Evidence-based education: Objections and future directions. *Frontiers in Education, 7*, Article 941410. https://doi.org/10.3389/feduc.2022.941410.

Didau, D. (2017, February 26). *Evidence and disadvantage: How useful is the EEF Toolkit?* Learning Spy. https://learningspy.co.uk/research/evidence-and-disadvantage/.

Di Pietro, G. (2023). The impact of Covid-19 on student achievement: Evidence from a recent meta-analysis. *Educational Research Review, 39*, Article 100530. https://doi.org/10.1016/j.edurev.2023.100530.

Driessen, G. (2017). The validity of educational disadvantage policy indicators. *Educational Policy Analysis and Strategic Research, 12*(2), 93–108. https://doi.org/10.5281/zenodo.6803502.

Driessen, G. (2021, March 23). *Doelmatig en effectief.* Didactief Online. https://www.researchgate.net/publication/350439317_Doelmatig_en_effectief.

Driessen, G. (2021, April 19). *NPO: Q&A's*. Didactief Online. https://doi.org/10.5281/zenodo.10902987.

Driessen, G. (2021, May 12). *Het NPO-menu: Gehakt.* Didactief Online. https://doi.org/10.5281/zenodo.10902976.

Driessen, G. (2021, June 30). *Driessen duikt in NPO.* Didactief Online. https://doi.org/10.5281/zenodo.10902957.

Driessen, G. (2022). *The many facets of educational disadvantage: Policies, interventions, effects*. Eliva Press. https://doi.org/10.5281/zenodo.6641161.

Driessen, G. (2025). The fragile foundation of pre- and early-school programs for disadvantaged children. *Forum for Education Studies, 3*(1), 1–16. https://doi.org/10.59400/fes1869.

EEF (2022a). *The Impact of COVID-19 on Learning: A review of the evidence*. London: Education Endowment Foundation. https://educationendowmentfoundation.org.uk/guidance-for-teachers/covid-19-resources/best-evidence-on-impact-of-covid-19-on-pupil-attainment

EEF (2022b). *Using the Toolkits—A summary*. https://d2tic4wvo1iusb.cloudfront.net/production/documents/toolkit/EEF-Toolkit-summary.pdf

EEF (2024). *EEF Annual Report September 2023–*

August 2024. https://d2tic4wvo1iusb.cloudfront.net/production/documents/annual-reports/eef-annual-report-2024-v1.0.0.pdf

EEF (2025a). *Teaching And Learning / Early Years Toolkit Guide*. https://d2tic4wvo1iusb.cloudfront.net/production/documents/toolkit/toolkit_guide_2025_v3.0.0.pdf

EEF (2025b). *Teaching And Learning Toolkit—Explainer.* https://d2tic4wvo1iusb.cloudfront.net/production/documents/toolkit/teaching_and_learning_toolkit_-_explainer_v.1.2.0.pdf

EEF (2025c). *EEF launches comprehensive updates to ten strands of the Teaching and Learning Toolkit*. https://educationendowmentfoundation.org.uk/news/updates-ten-strands-toolkit.

EEF (2026). *The Education Endowment Foundation*. https://educationendowmentfoundation.org.uk/.

Farquharson, C., Sibieta, L., & Waltmann, B. (2021). *COVID-related spending on education in England.* The Institute for Fiscal Studies. https://ifs.org.uk/publications/covid-related-spending-education-england-0.

François, M., & De Witte, K. (2025). *Addressing learning deficits and student well-being: International insights for the evaluation of the Dutch National Education Programme.* Katholieke Universiteit Leuven. https://www.rijksoverheid.nl/documenten/kamerstukken/2025/12/09/de-covid-19-pandemie-in-internationaal-perspectief-pdf.

Gambi, L., & De Witte, K. (2026). Academic resilience and policy reform in Belgium. In L. Volante, K. De Witte, O. Giancola, L. Salmieri, & D. A. Klinger (Eds.), *Risk, resilience, and recovery across global education systems: The long-term impact of pandemic reforms* (pp. 129-144). Routledge. https://doi.org/10.4324/9781003608806-10.

Giancola, O., & Salmieri, L. (2026). Academic resilience and policy reform in Italy. In L. Volante, K. De Witte, O. Giancola, K. Salmieri, & D. Klinger (Eds.), *Risk, resilience, and recovery across global education systems: The long-term impact of pandemic reforms* (pp. 70-89). Routledge. https://doi.org/10.4324/9781003608806-7.

Gilleece, L., & Clerkin, A. (2025). Towards more robust evaluation of policies and programmes in education: Identifying challenges in evaluating DEIS and Reading Recovery. *Irish Educational Studies, 44*(3), 445–473. https://doi.org/10.1080/03323315.2024.2334704.

Gorard, S., & Chen, W. (2025). How to assist research-informed education? *International Journal of Research and Innovation in Social Science, 9*(10), 4110–4124. https://doi.org/10.47772/IJRISS.2025.910000338.

Gorard, S., See, B. H., & Siddiqui, N. (2020). What is the evidence on the best way to get evidence into use in education? *Review of Education, 8*(2), 570-610. https://doi.org/10.1002/rev3.3200.

Haelermans, C., & Havermans, W. (2026). Academic resilience and policy reform in the Netherlands. In L. Volante, K. De Witte, O. Giancola, L. Salmieri, & D. Klinger (Eds.), *Risk, resilience, and recovery across global education systems: The long-term impact of pandemic reforms* (pp. 110-128). Routledge. https://doi.org/10.4324/9781003608806-9.

Haelermans, C., Huijgen, T., Jacobs, M., Levels, M., Van der Velden, R., Van Vugt, L., & Van Wetten, S. (2020). Using data to advance educational research, policy, and practice: Design, content, and research potential of the Netherlands Cohort Study on Education. *European Sociological Review, 36*(4), 643–662. https://doi.org/10.1093/esr/jcaa027.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* Routledge. https://inspirasifoundation.org/wp-content/uploads/2020/05/John-Hattie-Visible-Learning_-A-synthesis-of-over-800-meta-analyses-relating-to-achievement-2008.pdf.

Heijsters, L., Van der Ploeg, S., & Weijers, S. (2020). *Schaalgrootte in het primair en voortgezet onderwijs: Een literatuurstudie*. Oberon. https://www.vosabb.nl/wp-content/uploads/2020/06/literatuurstudie-schaalgrootte-in-het-primair-en-voortgezet-onderwijs.pdf.

Higgins, S., Villanueva Aguilera, B., Dobson, E., Gascoine, L., Kalambouka, A., Katsipataki, M., Reardon, J., Rovida, M., & Uwimpuhwe, G. (2024). *EEF evidence database Phase 2: Protocol and analysis plan*. Education Endowment Foundation. https://d2tic4wvo1iusb.cloudfront.net/production/documents/toolkit/EEF_Database_Backfilling_Protocol_v05_Jan24.pdf.

Jones, G. (2019a, May 31). *"It's time we changed" converting effect sizes to months of learning is seriously flawed.* https://www.garyrjones.com/blog/2019/5/31/its-time-we-changed-converting-effect-sizes-to-months-of-learning-is-seriously-flawed..

Jones, G. (2019b, September 5). *The myth of a 0.4 SD effect size and a year's worth of progress*. https://www.garyrjones.com/blog/2019/9/5/04-sd-and-a-years-worth-of-a-progress.

Katsipataki, M., & Higgins, S. (2016). What works or what's worked? Evidence from education in the United Kingdom. *Procedia - Social and Behavioral Sciences, 217*, 903–909. https://doi.org/10.1016/j.sbspro.2016.02.030.

Kraft, M. (2020). Interpreting effect sizes of education interventions. Educational Researcher, 49(4), 241–253. https://doi.org/10.3102/0013189X20912798.

Luyten, H., Merrell, C., & Tymms, P. (2017). The contribution of schooling to learning gains of pupils in Years 1 to 6. *School Effectiveness and School Improvement, 28*(3), 374–405. https://doi.org/10.1080/09243453.2017.1297312.

Maurya, P., & Moulin, L. (2026). Academic resilience and policy reform in France. In L. Volante, K. De

Witte, O. Giancola, K. Salmieri, & D. Klinger (Eds.), *Risk, resilience, and recovery across global education systems: The long-term impact of pandemic reforms* (pp. 30-50). Routledge. https://doi.org/10.4324/9781003608806-5.

Ministry of Education, Culture and Science. (2021a). *National Education Programme: Support programme forrecoveryandperspective.*https://www.rijksoverheid.nl/documenten/kamerstukken/2021/02/17/nationaal-programma-onderwijs-steunprogramma-voor-herstel-en-perspectief.

Ministry of Education, Culture and Science. (2021b). *Annex to the National Education Programme: Support programme for recovery and perspective*. https://open.overheid.nl/documenten/ronl-d09b8087-fe33-4349-93ef-af1bf5af822f/pdf.

Ministry of Education, Culture and Science. (2021c). *Further elaboration of the National Education Programme*. https://www.rijksoverheid.nl/documenten/kamerstukken/2021/05/21/nadere-uitwerking-nationaal-programma-onderwijs.

Ministry of Education, Culture and Science. (2022, May 23). *Regulation of the Minister for Primary and Secondary Education of 16 May 2022, no. 32561416.* Staatscourant. https://zoek.officielebekendmakingen.nl/stcrt-2022-13905.pdf.

Ministry of Education, Culture and Science. (2025a). *From crisis to opportunity: Final report on the National Education Programme in primary and secondary education.* https://www.rijksoverheid.nl/documenten/kamerstukken/2025/12/09/ocw-eindrapportage-np-onderwijs.

Ministry of Education, Culture and Science. (2025b). *School careers: NP Education 2025*. https://duo.nl/open_onderwijsdata/images/schoolloopbanen-np-onderwijs-2025.pdf.

Ministry of Education, Culture and Science. (2025c). *School careers 2025: Technical appendix.* https://duo.nl/open_onderwijsdata/images/technische-bijlage-schoolloopbanen-np-onderwijs-2025.pdf.

NJi. (2026, March 12). Effectgrootte. https://www.nji.nl/kennis/effectieve-jeugdhulp/effectgrootte.

NKO. (2026, March 12). *Toolkits manual. National Knowledge Institute for Education.* https://www.onderwijskennis.nl/handleiding-toolkits.

Peeters, W. (2025). *Evidence-informed & the complexity of practice.* https://vernieuwenderwijs.nl/evidence-informed-praktijk/.

RED-team Onderwijs. (2021, March 25). *Red card for the National Education Programme* (primary and secondary education). https://redhetonderwijs.com/rode-kaart-voor-nationaal-programma-onderwijs-po-en-vo/.

Reimers, F., & Schleicher, A. (2020). *A framework to guide an education response to the COVID-19 pandemic of 2020*. OECD. https://doi.org/10.1787/6ae21003-en.

Scheerens, J., & Kirschner, P. (2021, May 21). *NPO menu: A scattergun approach*. Didactief Online. https://didactiefonline.nl/blog/paul-kirschner/menukaart-npo-een-hagelschotbenadering.

Schnepf, S., Volante, L., Klinger, D., Giancola, O., & Salmieri, L. (Eds.) (2024). *The pandemic, socioeconomic disadvantage and learning outcomes: Cross-national impact analyses of education policy reforms*. Publications Office of the European Union. https://doi.org/10.2760/800165.

Schult, J., Kürzinger, A., & Schneider, R. (2026). Academic resilience and policy reform in Germany. In L. Volante, K. De Witte, O. Giancola, K. Salmieri, & D. Klinger (Eds.), *Risk, resilience, and recovery across global education systems: The long-term impact of pandemic reforms* (pp. 51-69). Routledge. https://doi.org/10.4324/9781003608806-6.

Simpson, A. (2017). The misdirection of public policy: comparing and combining standardised effect sizes. *Journal of Education Policy, 32*(4), 450–466. https://doi.org/10.1080/02680939.2017.1280183.

Simpson, A. (2018). Princesses are bigger than elephants: Effect size as a category error in evidence-based education. *British Educational Research Journal, 44*(5), 897–913. https://doi.org/10.1002/berj.3474.

Slavin, R., & Smith, D. (2009). The relationship between sample sizes and effect sizes in systematic review in education. *Educational Evaluation and Policy Analysis, 31*(4), 500–506. https://doi.org/10.3102/0162373709352369.

Styles, B., & Torgerson, C. (2018). Randomised controlled trials (RCTs) in education research –Methodological debates, questions, challenges. *Educational Research, 60*(3), 255–264. https://doi.org/10.1080/00131881.2018.1500194.

Tweede Kamer (2023). *Letter from the Minister for Primary and Secondary Education* (Kamerstuk 31293, No. 710). https://zoek.officielebekendmakingen.nl/kst-31293-710.pdf.

Tweede Kamer (2025). *Letter from the State Secretary of Education, Culture and Science* (Kamerstuk 31293, No. 858). https://zoek.officielebekendmakingen.nl/kst-31293-858.pdf.

Van den Broek, A., Termorshuizen, T., Luyten, E., Cuppen, J., & Groot, A. (2025). *Evaluation of NP Education. ResearchNed*. https://www.researchned.nl/publicatie/evaluatie-nationaal-programma-onderwijs-funderend-onderwijs/.

Van der Steeg, M. (2025). Learning delays caused by COVID-19 have largely been recovered. *ESB, 110*(4828), 345–347. https://esb.nu/door-corona-ontstane-leervertragingen-zijn-grotendeels-hersteld/.

Van Vugt, L., Haelermans, C., Baumann, S., & Ronda, S. (2025a). *Learning growth and COVID-19 factsheets: Autumn 2025 – Technical explanation*. https://www.nationaalcohortonderzoek.nl/toelichting-factsheets.

Van Vugt, L., Haelermans, C., Baumann, S., Hendrikse, A., & Ronda, S. (2025b). *Five years since COVID-19:*

Similar learning growth to the pre-COVID-19 period (NCO Factsheet No. 2025-1). https://www.nationaalcohortonderzoek.nl/factsheets-leergroei.

Van Vugt, L., Haelermans, C., Baumann, S., Hendrikse, A., & Ronda, S. (2025c). *Five years since COVID-19: By parental education level and school weighting* (NCO Factsheet No. 2025-2). https://www.nationaalcohortonderzoek.nl/factsheets-leergroei.

Volente, L., & Klinger, D. (2026). The global impact of the pandemic on student outcomes. In Volente et al. (Eds), *Risk, resilience, and recovery across global education systems*: *The long-term impact of pandemic reforms* (pp. 3-12). Routledge. https://doi.org/10.4324/9781003608806-2.

Volante, L., De Witte, K., Giancola, O., Salmieri, K., & Klinger, D. (Eds.) (2026). *Risk, resilience, and recovery across global education systems*. The long-term impact of pandemic reforms. Routledge. https://doi.org/10.4324/9781003608806.

Wikström, C., & Wikström, M. (2026). Academic resilience and policy reform in Sweden. In L. Volante, K. De Witte, O. Giancola, K. Salmieri, & D. Klinger (Eds.), *Risk, resilience, and recovery across global education systems: The long-term impact of pandemic reforms* (pp. 90-109). Routledge. https://doi.org/10.4324/9781003608806-8.

Wrigley, T. (2016). Not so simple: The problem with "evidence-based practice" and the EEF toolkit. *FORUM, 58*(2), 237–250. https://doi.org/10.15730/forum.2016.58.2.237.

Wrigley, T. (2018). The power of "evidence": Reliable science or a set of blunt tools? *British Educational Research Journal, 44*(3) 359–376. https://doi.org/10.1002/berj.3338.