

A COVID-19 Education Recovery Program

Subjects: [Education & Educational Research](#)

Contributor: Geert Driessen

As a consequence of the COVID-19 pandemic, many students have developed substantial educational delays, both cognitively and social-emotionally. To counter such negative effects of the school closures, several policies and support strategies on attainment and social-emotional well-being have been proposed and implemented. In the Netherlands, the focus is on using evidence-based interventions to boost educational achievement. The question is, however, how evidence-based the interventions really are.

COVID-19

educational delays

recovery program

the netherlands

effect sizes

evidence-based

interventions

school closures

learning loss

educational inequality

1. National Program Education

As a consequence of the COVID-19 pandemic, many students, from pre-primary to university education institutions, have developed substantial educational delays, both cognitively and social-emotionally, and especially students from low socioeconomic and ethnic/immigrant backgrounds.^{[1][2][3][4]} To counter the negative effects of the school closures, several policies and support strategies on attainment and social-emotional well-being have been proposed and implemented.^{[5][6]}

In The Netherlands, the Ministry of Education has introduced an unprecedented National Program Education (NPE) to combat Covid-19 related educational delays.^[7] The core of this program consists of providing schools with additional budgets which they can spend on implementing evidence-based interventions. The total budget amounts to €8.5 billion for a 2.5 year period. Primary and secondary schools receive at least €700 per student per year; depending on the social-ethnic backgrounds of their students, schools may receive even more. Schools are required to spend the money on evidence-based interventions, denoting approaches that emphasize the practical application of the findings of the best available current scientific research, preferably from experiments using randomized controlled designs. The Ministry has developed a so-called menu card with (for now) 22 interventions schools can choose from.^[8] Most of them (19) come from the Teaching and Learning Toolkit developed by the English Education Endowment Foundation (EEF);^[9] the rest (3) is added by the Dutch Ministry.

2. Timeline

For 2021, the Ministry has published a timeline for the schools, depicted in Table 1.^[10] This is a very tight one, as the schools must do a lot of work in a very short period of time (and while at the same time reopening the schools

after a lockdown) and - inevitable - partly during the summer holidays.

Table 1. Timeline National Program Education (February 2021 – December 2021)

February	Ministry announces National Program Education.
March	Schools receive a letter with information on the contents of the program.
April	Schools make a so-called school scan of specific problems and needs of the students and school. The focus is on the students' cognitive development (all domains, not only language and math) and their social-emotional development and well-being. Information comes from available student monitoring systems, tests, observations, questionnaires and documents. Missing information can be supplemented by new data collections. All this input is then analyzed and results in an up-to-date overview of problems and needs.
May	Based on the school scan, schools choose from the menu card with evidence-based interventions.
June	Schools are informed as to how large their NPE budget will be.
July - August	Schools write a school program with all of their measures for the school years 2021/22 and 2022/2023. This program must be approved of by the Participation Council.
September	Schools receive their NPE budget and start with the implementation of the school program.
End of the year	In their annual report school boards provide a summary of their NPE approach and results.

3. Effect sizes

Most of the evidence-based interventions on the NPE menu card are translations of the EEF's Teaching and Learning Toolkit.^[9] The Toolkit's interventions specifically aim at boosting the attainment of disadvantaged children. In the Toolkit's Technical Appendices the effectiveness of each of the interventions is accounted for. EEF uses three criteria to rate the interventions: (1) the cost; (2) the strength of the evidence; (3) the impact in terms of months additional progress, where it is assumed that 1 SD is equivalent to 1 year of progress, and thus 0.09 SD is about 1 month of progress^[11] (a procedure which is severely criticized^{[12][13][14][15]}).

The interventions' effects are expressed by EEF in terms of so-called effect sizes. There are many types of effect sizes, but in its simplest form an effect size represents the standardized difference between two groups, ideally the experimental group and the control group (or the group with the intervention and the group without the intervention, respectively). Because they are standardized, they can be compared across studies. A common effect size is Cohen's *d*. Cohen (1988)^[16] himself provided the following rule of thumb for interpreting *d*: <0.20, negligible effect; 0.20-0.50, small effect; 0.50-0.80, medium effect; >0.80, large effect. However, he cautions that these values are necessarily somewhat arbitrary and acknowledges that the qualifications should be used in the proper context. According to Hattie (2009)^[17], who performed many meta-meta analyses using *d*, most interventions in education 'work' with an average effect size of about 0.40, and that this should therefore be a benchmark. EEF (2018)^[11] uses the following classification: <0.02, negative, no or very low effect; 0.02-0.18, low effect; 0.19-0.44, moderate effect; 0.45-0.69, high effect; 0.70-1.00, very high effect. On the one hand, EEF argues that a 'small' effect may nevertheless be educationally important if it is easy or cheap to attain or achievable with groups who are otherwise hard to influence. On the other hand, a 'large' effect size may not be important if it is unrealistic to bring about in normal circumstances. One of the contributors to the NPE menu card is the Netherlands Youth Institute (NJI, 2021),^[18] which employs Cohen's rule of thumb interpretation: <0.20, no or a negligible effect; 0.20- 0.50, a small effect; 0.50-0.80, a medium effect; >0.80, a large or very large effect. But NJI, too, points to the fact that one should also consider the context when interpreting effect sizes.

4. Evidence-based interventions?

The main question here is how much evidence there really is for the evidence-based interventions included in the menu card. To get an impression of the effectiveness, the evidence supplied by EEF in the Toolkit's Technical Appendices^[9] has been critically reviewed and summarized. EEF discerns two types of studies, single studies and meta-analyses (which comprise a quantitative summary of a series single studies). Table 2 provides per intervention the number of underlying meta-analyses and single studies, the 'range' (minimum – maximum) of the effect sizes (*d*), the average effect, and prominent features as regards content and methodology.

Table 2. Overview effects evidence based interventions (effect sizes *d*; source: EEF Toolkit)

Intervention	Number of meta-analyses / single studies	Minimum – maximum effect	Average effect	Comments
Early years interventions	12 / 0	0.15 – 0.55	0.38*	Methodological problems. More recent studies → weaker effects. Mainly older US studies; UK studies → no consistent effects.
Extending school time	8 / 5	-0.14 – 0.40	0.11*	Substantial variation in effects, not consistent. Mainly US studies; UK studies → little evidence. Often part of broader program → origin effect not clear. Problem: absenteeism students.
Summer schools	6 / 4	0.00 – 0.43	0.18*	Small groups, intensive, experienced teachers → more successful. Problem: absenteeism students.
One to one tuition	7 / 7	-0.06 – 0.70	0.37*	Sometimes schooling tutor important, sometimes not. Substantial variation in effects, not consistent.
Individualized instruction	7 / 2	-0.07 – 0.41	0.19**	Many older studies. Methodological problems. Much variation in effects.
Small group instruction	4 / 4	-0.08 – 1.61	0.31*	Mainly older studies. Mainly secondary education. Not aimed at improving achievement. Quality instruction probably important. Enormous variation in effects, not consistent.
Direct instruction#	?	?	?	

Intervention	Number of meta-analyses / single studies	Minimum – maximum effect	Average effect	Comments
Peer tutoring	9 / 4	-0.06 – 1.05	0.37*	Some older studies. Methodological problems. Substantial variation in effects, not consistent. More recent studies → stronger effects, often in combination with digital technologies.
Feedback	7 / 0	0.20 – 0.97	0.63*	Nearly all (very) old studies. Much variation in effects. More recent studies → weaker effects.
Mastery learning	6 / 4	0.04 – 1.64	0.40*	Mainly very old studies. Methodological problems. Enormous variation in effects, not consistent. More recent UK studies → weaker effects.
Reading comprehension techniques	8 / 4	0.10 – 0.74	0.45*	Mainly US studies. More recent UK studies → weaker effects.
Oral language interventions	11 / 7	-0.14 – 0.91	0.37*	Very broad concept with many interpretations. Enormous variation in effects, not consistent.
Well-being interventions#	?	?	?	
Sports participation	3 / 0	0.10 – 0.80	0.17**	Methodological problems. Substantial variation in effects, not consistent.
Arts participation	5 / 2	0.03 – 0.77	0.15*	Methodological problems. Substantial variation in effects, not consistent.

Intervention	Number of meta-analyses / single studies	Minimum – maximum effect	Average effect	Comments
Meta-cognition and self-regulation	11 / 9	-0.14 – 0.90	0.54*	Effect strongly depends on whether students are able to take responsibility for their learning process. Methodological problems. Enormous variation in effects, not consistent. More recent studies → weaker effects.
Collaborative learning	11 / 1	0.13 – 0.91	0.38*	Substantial variation in effects, not consistent.
Reducing class size	6 / 0	0.12 – 0.34	0.19**	Mainly (very) old US studies. Only (small) effects in case of less than 20 or even 15 students per class.
Teaching assistants	0 / 15	-0.15 – 1.50	0.08***	Enormous variation in effects, not consistent.
Supportive conditions#	?	?	?	
Parental engagement	15 / 4	-0.14 – 0.65	0.22*	Very broad concept with many interpretations. Mainly older US studies.
Digital technology	32 / 2	-0.15 – 1.13	0.29*	Very broad concept with many interpretations. Methodological problems. Enormous variation in effects, not consistent.

Not part of the EEF Toolkit; * Weighted mean; ** Median; *** Indicative.

When looking at the average effect sizes in Table 2 and focusing on the 19 EEF interventions (the other 3 are added by the Dutch Ministry of Education), it can be concluded that according to Cohen's (1988)^[16] rule of thumb, 37 percent of the interventions have a negligible effect, 53 percent a small effect, and only 2 a medium effect. According to Hattie (2009)^[17], only 4 out of 19 interventions, or 21 percent, achieve at or above his benchmark level of 0.40. It is obvious that EEF's^[11] rule of thumb is the most lenient one: 26 percent of the interventions have a low effect, 58 percent a moderate effect, and 16 percent a high effect. To summarize these findings: in so far interventions are effective, only very few of them really show a robust effect. In addition, from a further look at the range of the effect sizes, it becomes clear that this conclusion is strongly embellished. There exists an enormous variation in the magnitude of the effect sizes, they vary from negative to strongly positive. (In fact, there is so much variation that computing averages is not a statistical sound procedure.) What is lacking in most of the cases is consistency. This variability probably is caused by the fact that under one intervention (e.g., parental engagement), there are many practical interpretations (e.g., helping with homework, being a member of the parent council, volunteering in the classroom, or visiting libraries), with each having its specific effects (e.g., negative or positive) on different output domains (e.g., language or motivation) for different groups of students (e.g., immigrant or high social class). Therefore, when head teachers choose an intervention to solve the specific problems and needs of their schools, they have no guarantee whatsoever that their local interpretations will lead to the hoped for results. The situation clearly is considerably more complex than assumed at first sight. In addition, there are many more problematic considerations with regard to the evidence base of the interventions. Table 3 provides an overview.

Table 3. – Overview problematic aspects studies into evidence based interventions (trends)

- Interventions include an abundance of (practical) interpretations; mostly very broad, abstract concepts; often not clear what precisely is meant.
- No or hardly any targeted research; as far there is research, it mainly is from the US and sometimes UK; Dutch studies are lacking.
- EEF conclusions are based on a relatively small number of studies, which, moreover, often are dated i.e. more than 20 years old.
- The EEF meta and single studies show considerable overlap, especially when conducted by the same authors.
- No studies from a Covid-19 context. Different goal studies and different target group: interventions aim preventing and combating educational delays caused by factors in the home situation of children from low socioeconomic and ethnic backgrounds.
- Research often does not meet methodological standards.

- Both positive and negative effects and also often null effects; from strongly positive to strongly negative; consistency and unequivocality are lacking.
- Positive effects seem to be over-reported; less or no attention to (causes of) negative and null effects.
- Effects only valid for specific target groups (e.g., socioeconomic background, ethnicity, age, grade, educational sector) in combination with subject related domains (e.g., language, math, self-reliance, motivation). In the summary effect sizes all such nuances are lumped together, while they are crucial for an appropriate application of interventions to combat Covid-19 delays.
- The interpretation of effects is not always in line with standards; effective is not effective according to the own standards.
- More recent studies often show weaker effects than older studies.
- Much discussion and doubt regarding the validity of transformation of effects into “additional months of progress”.

5. Conclusions

On the basis of these findings, the obligation for schools to choose so-called evidence-based interventions from the NPE's menu card is a highly dubious one. First of all, it appears that many interventions are really not that evidence-based as one would expect. In addition, it cannot be expected from the schools' head teachers to make a valid and reliable choice as they simply do not have the statistical and methodological knowledge to interpret the EEF's scientific foundation of the interventions correctly. Moreover, the evidence presented in the EEF's Toolkit only holds for a specific interpretation of an intervention, for specific groups (e.g., age, socioeconomic background, ethnicity) and domain (e.g., language, motivation). Probably it is a better idea to now start well-designed studies, preferably randomized controlled experiments, to have really proven interventions available in the future. The present situation, with all that money and data available under the Dutch National Program Education, offers an ideal starting point.

References

1. Per Engzell; Arun Frey; Mark D. Verhagen; Learning loss due to school closures during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences* **2021**, *118*, 1-7, 10.1073/pnas.2022376118.

2. Best evidence on impact of Covid-19 on pupil attainment . Education Endowment Foundation. Retrieved 2021-6-23
3. Grewenig, E., et al.. COVID-19 and Educational Inequality: How School Closures Affect Low- and High-Achieving Students; Institute of Labor Economics: Bonn, Germany, 2020; pp. 1-31.
4. Bailey, D., et al.. Achievement Gaps in the Wake of COVID-19; Annenberg Institute, Brown University: Providence, RI, 2021; pp. 1-34.
5. Reimers, M.; Schleicher, A.. A framework to guide an education response to the COVID-19 pandemic of 2020; OECD: Paris, 2020; pp. 1-40.
6. Rose, S., et al.. Impact of school closures and subsequent support strategies on attainment and socio-emotional wellbeing in key stage 1: Interim paper 1; Education Endowment Foundation, National Foundation for Educational Research: London, 2021; pp. 1-14.
7. Nationaal Programma Onderwijs . Ministerie van Onderwijs, Cultuur en Wetenschap. Retrieved 2021-6-23
8. Menukaart interventies funderend onderwijs . Ministerie van Onderwijs, Cultuur en Wetenschap. Retrieved 2021-6-23
9. Teaching and Learning Toolkit . Education Endowment Foundation. Retrieved 2021-6-23
10. Tijdljn Nationaal Programma Onderwijs . Ministerie van Onderwijs, Cultuur en Wetenschap. Retrieved 2021-6-23
11. Higgins, S., et al.. Sutton Trust-EEF Teaching and Learning Toolkit & EEF Early Years Toolkit. Technical appendix and process manual; Education Endowment Foundation: London, 2018; pp. 1-56.
12. Matthew D. Baird; John F. Pane; Translating Standardized Effects of Education Programs Into More Interpretable Metrics. *Educational Researcher* **2019**, *48*, 217-228, 10.3102/0013189x19848729.
13. Dadey, N.; Briggs, D.; A Meta-Analysis of Growth Trends from Vertically Scaled Assessments. *Practical Assessment, Research & Evaluation* **2012**, *17*, 1-15.
14. Adrian Simpson; Princesses are bigger than elephants: Effect size as a category error in evidence-based education. *British Educational Research Journal* **2018**, *44*, 897-913, 10.1002/berj.3474.
15. It's time we changed. Converting effect sizes to months of learning is seriously flawed . Gary Jones. Retrieved 2021-6-23
16. Cohen, J. . Statistical Power Analysis for the Behavioral Sciences; Academic Press: New York, NY, 1988; pp. 1-473.

17. Hattie, J.. Visible Learning : A Synthesis of Over 800 Meta-Analyses Relating to Achievement; Routledge: Oxon, 2009; pp. 1-378.

18. Effectgrootte . Nederlands Jeugdinstituut. Retrieved 2021-6-23

Retrieved from <https://encyclopedia.pub/entry/history/show/26533>