Forum for Education Studies 2026, 4(3), 4858.
https://doi.org/10.59400/fes4858

Review

Why the educational disadvantage policy fails. A case study of the Netherlands

Geert Driessen

Radboud University, 6525XZ Nijmegen, Netherlands
*** Corresponding author:** Geert Driessen, driessenresearch@gmail.com

CITATION

Driessen, G. Why the educational disadvantage policy fails. A case study of the Netherlands. Forum for Education Studies. 2026; 4(3): 4858. https://doi.org/10.59400/fes4858

COPYRIGHT

Copyright © 2026 by author(s). *Forum for Education Studies* is published by Academic Publishing Pte Ltd. This work is licensed under the Creative Commons Attribution (CC BY) license. https://creativecommons.org/licenses/by/4.0/

Abstract: For over half a century, the Dutch Ministry of Education has been implementing policies aimed at combating educational disadvantages linked to the socio-economic circumstances in which children grow up. The focus is on all children who meet the relevant disadvantage criteria, whether they have developmental delays or not. This policy has by now cost around forty billion euros. Unfortunately, student performance continues to decline. This in-depth literature review describes the manner in which those billions are disbursed to school boards and municipalities, and the way in which they are deployed. In view of this failure of a group-oriented approach, a plea is made for an individual-oriented approach. In contrast to the preventive group approach, the individual-centered approach focuses on the actual developmental delays of individual children and is more remedial in nature.

Keywords: educational disadvantage policy; evidence-based education; interventions; evaluation; effectiveness; case-study; the Netherlands

1. Introduction

More than fifty years ago, in 1974, the Dutch Ministry of Education initiated the Educational Stimulation Policy, which was aimed at improving the educational opportunities of (native) working-class children. In addition, there was the Cultural Minority Policy, for children of migrant workers and from former colonies, as well as a policy for girls. In 1985, all of this was integrated into the Education Priority Policy, which was later renamed the Educational Disadvantage Policy (EDP).

In this article, I will examine in succession (1) exactly what the target group of the EDP is, (2) to which measures the tens of billions paid out by the government to school boards and municipalities were spent, and (3) what those investments have yielded, or in other words: has the central objective, namely the prevention and combating of educational disadvantages resulting from socio-ethnic factors related to the home situation, been achieved? Finally, I draw some conclusions and present suggestions for possible alternatives.

2. The target group

2.1. The individual approach

Politically, opinions regarding the financing of these policy tracks differed from the outset (Driessen, 2015). The Right (Liberal Party) favored an individual approach; the origin of children's learning delays lies with the child itself. Extra support should therefore be aimed at those children who truly need it. This could be achieved by identifying deficits through observations and assessments and intervening as needed. It is logical to start doing this at an early age. However, at the time, there was a strong aversion to the formalization of preschool education and certainly to the testing

of young children (in 1985, Dutch preschool and primary school were merged): preschoolers were supposed to play, not learn, was the opinion of many (former) preschool teachers.

Since the 2014/15 school year, schools have been required to monitor the language and math development of their students using a Student Monitoring System (SMS). The education sector has mixed feelings about this. Research (Geurtsen, 2023) shows that it places pressure on both students and teachers, and leads to stress for some of the students. It also appears that many students practice for the tests outside of school, which raises the question of what is actually being tested. Goorhuis (2026) argues that young children cannot be reliably tested. This is because their development is erratic and leapfrog, sometimes showing progress, sometimes stagnation, and sometimes even relapse. From her own research, she concluded that children can only be reliably tested at the age of 6. In this context, attention is drawn to the weak predictive validity of tests administered to young children (Leseman, 2004): the level of their test scores says little about the level of later school performance. Initially, the obligation to administer such tests applied to all grades (i.e. 1-8) in primary education. However, as of the 2022/23 school year, due to the aforementioned objections, the Ministry no longer permits SMS tests to be administered in kindergarten grades 1 and 2. It is, however, permitted to use observation instruments there. This immediately sparked debate. While it was initially stated that schools are free to choose an observation instrument, the minister later indicated that only approved instruments were permitted—that is to say, those that are validated, have sufficient reliability, and are (absolutely) standardized. This led some to lament that the banned kindergarten tests have returned in a new form (Van den Burg-Poortvliet, 2021).

2.2. The group approach

Opposite the individual perspective of the right stood the group-based approach of the left (Labor Party). Right from the start, this has been the guiding principle of the official approach to educational disadvantages. The origin of these disadvantages lies in ('deficiencies' in) the environment in which the children in question grow up. Subsidies must be made available for all children growing up under such 'unfavorable' circumstances. Specifically, schools (and municipalities, including welfare institutions) were to receive subsidies for both native and immigrant working-class children, in order to work on preventing and combating socio-economically determined educational disadvantages. An advantage of the group-based approach is, certainly when started at an early stage, that the emphasis lies on prevention; this is in contrast to the individual approach, which is aimed more at remediation. The first approach is based on expected school performance, the second on actual performance.

Initially, the allocation of the subsidy took place via the parental 'occupational factor' (manual labor occupation or not); later, this became the 'student weight' (with the indicators being education, occupation, and ethnicity of the students' parents), and even later, only parental education level as an indicator (Ladd & Fiske, 2009). The advantage of this approach was that the required data were logical and transparent, and that schools also knew for whom they were receiving the subsidy. A disadvantage was that it entailed quite a bit of bureaucratic red tape: the schools had to collect the necessary data themselves. Partly as a result of this, the system was susceptible to fraud, which is why checks were carried out by the Inspectorate of Education.

The biggest problem, however, was the questionable accuracy of the allocation criteria. The three indicators of student weight (education, occupation, ethnicity)

initially (in 1985) explained about 25 percent of the differences in educational performance; the so-called predictive validity was therefore 0.50. After occupation and parents' ethnicity were removed, and only education remained as an indicator of disadvantage, a predictive validity of 0.20 remained, and consequently no more than 4 percent of variance explained (which is also reported in international studies; Liu et al., 2022; Sirin, 2005; Zhao et al., 2025). The latter means that 96 percent of performance differences between students are related to factors other than the level of education attained by their parents (Driessen, 2017). Consequently, there was only a weak basis for distributing billions of euros in subsidies, with a very high number of false positives and false negatives, or in other words, a very large number of children for whom a school wrongly received a subsidy or, respectively, was denied a subsidy. According to Leseman (2004), the weighting system is insufficiently specific (extra money for students who do not need it) and insufficiently sensitive (no extra money for students who do need extra support, but who are not considered belonging to disadvantaged groups on socio-economic or ethnic grounds).

In order to address these disadvantages, Statistics Netherlands (Posthumus et al., 2018) designed a new scheme with six risk indicators: father's education, mother's education, parents' origin, mother's length of residence, parents in debt restructuring, and average educational level of mothers per school. These six characteristics explained 19.8 percent of the performance differences. It should be noted that two indicators were exceptionally skewed, meaning that there were only a few families to whom they applied. These concerned the length of residence of the mothers and parents in debt restructuring, for which the correlation with the test scores was also negligibly weak ($r = 0.03$ and $r = 0.08$, respectively). In addition, the correlation between the average educational level of the mothers at a school and the test scores was not significant. The advantage of this new system was that schools no longer had to collect this information themselves and were thereby relieved of this burden. After all, these data were located in databases available to Statistics Netherlands (only). Another positive point was that the predictive validity was significantly higher than that of the previous scheme, which was based solely on parental education level. However, it was still somewhat lower than that of the original scheme involving education, occupation, and ethnicity. On the other hand, there are several disadvantages. With three dubious indicators that are not significant or have only a very weak relationship with test performance, it is questionable why they were included at all, and why the ‘original’ indicators of education and origin did not suffice. Moreover, there is another important disadvantage to practical application. Statistics Netherlands is the only body that possesses all the data and calculates a total score based on this (meaningless to schools), which is then used to distribute the government subsidy among the schools. Previously, schools knew which of their students were disadvantaged children and for which they therefore received subsidies. Because all data are now stored on servers at Statistics Netherlands, schools are no longer aware of this. Through their governing bodies and as part of the lump sum, they simply receive a (large) bag of money with which they can, in principle, do whatever they want. The relationship between the awarded subsidy and the target group student has completely disappeared. A complete depersonalization, in other words.

Recently, Statistics Netherlands (CBS, 2026) revised its previous scheme. The core of the methodology remained the same. Using initially some 220 variables, the researchers attempted to determine which of these could best explain differences

in language and math achievement in grade 8 (the last year of primary education). With a selection of 10 characteristics (59 variables), between 13 and 16% of the differences in test performance could be explained. This is therefore less than under the currently applicable scheme (namely 19.8%). The characteristics are: father's education, mother's education, household income, household in problematic debts, parents' ethnic origin, family structure, length of stay of mother in the Netherlands, parents suspected of a crime, socio-economic status of the school population, and subpopulation regarding parents' educational level. Is that sufficient to distribute many billions in subsidies based on this? That is not common practice. The Education Endowment Foundation (EEF; Hunt et al., 2026), for example, does not consider 30 percent explained variance to be much. Based on an analysis of Dutch articles in this field published between 1975 and 1998, Kalmijn and Kraaykamp (1999) arrive at an average of 27 percent explained variance, although there is a substantial spread around that average. Similarly, Smedslund et al. (2022) examined psychological articles published between 1956 and 2022 and arrived at an average of 42.8 percent explained variance. Of course, one should not look solely at the proportion of explained variance, but the above is certainly food for thought. After all, it concerns a great many children and a great deal of money, year after year. The aforementioned disadvantages, incidentally, remain. Even now, schools do not know for which children they receive subsidies and to whom they should actually spend them. This can be qualified as a particularly peculiar state of affairs. And it invites inefficient and ineffective policy. In addition, some variables are probably extremely skewed. This seems, at least, to apply to ‘suspected of a crime’ (0.8% of 12-year-olds and older nationwide). There is another factor at play here. Someone who is suspected is not necessarily guilty, but the child is still branded (Kloprogge, 2026). Moreover, schools receiving substantial subsidies now run a high risk of being associated with a student body of criminal parents with large debts.

3. The allocation of subsidies

Below, I will first discuss a general measure, the weighting scheme, intended for both native and non-native disadvantaged children, and subsequently, as examples, two specific measures, namely Mother Tongue Education (MTE) for non-Dutch speaking children, and Early Childhood Education and Care programs (ECEC), for young native and non-native disadvantaged children.

3.1. The weighting scheme

A huge amount of money is involved in the EDP, and has been since 1974. It has become a whole industry, involving not only public parties but also an increasing number of commercial businesses. It is difficult to ascertain exactly which budgets are involved. Over all those years, money has been provided from various ministries (particularly Education, and Welfare) to diverse and changing recipients, mostly primarily school boards and municipalities (Onderwijsraad, 1998). By far the largest amount of money is contained in ‘the weighting scheme’, where children are assessed according to their degree of disadvantage and primary school boards receive subsidies on the basis of this. However, for a number of years now, ECEC (educational activities for children aged 2.5 to 6 from disadvantaged backgrounds) has also been funded in this manner. The same applies to the billions disbursed to schools in recent years within the framework of the National Education Program (NEP; intended to address COVID-19 learning gaps). In addition, collaborative projects at the regional level were funded by the EDP (such as in the Education Priority Areas), area-based

subsidies were provided (Impulse Areas), large-scale national measures were funded, such as the MTE, Dutch as a Second Language, and Intercultural Education (Driessen, 2022), various large 'standalone' projects falling under the EDP have been initiated, and municipalities can also use own funds to combat educational disadvantages. All in all, a considerable sum; I estimate it to be close to forty billion euros by now (for a partial overview, see Tweede Kamer, 2000).

But what has happened to all that money? It is hard to say. The administration in this area is sparse and, in any case, general and vague, and moreover, subject to continuous change. The tens of billions from the weighting scheme have largely ended up with the school boards as part of the lump sum. The latter implies that the boards are, in principle, free to spend it. Although the subsidy was awarded based on individual, personal characteristics (the child, parents, and school), there is no obligation for the recipients (schools, municipalities) to use the money specifically for these children. The subsidy is not earmarked. This could mean, for example, that school boards pool the resources of all schools under their authority and concentrate their use on one or a few schools; the other schools (and their disadvantaged students) then receive nothing. The subsidy can even be used for matters other than combating educational disadvantage.

While there has been an accountability obligation to the Inspectorate of Education for some time, this is very general: the school boards are required to mention the expenditure in their annual report, but it is anything but specific. Moreover, it is unclear to what extent school leaders and teachers are involved in the distribution of the EDP funds and are aware of who receives what for what purpose. An early evaluation of the EDP (Mulder, 1996) showed that at the time not only were there numerous primary schools that were unaware that they had been receiving extra subsidies for many years, but also that only a small proportion of those schools actually used these funds for specific activities aimed at combating disadvantage. The extra funds were generally considered part of the standard basic staffing, which in practice usually led to smaller class sizes. This information comes from an older study, but more recent research has shown that little change has occurred in subsequent years.

The Court of Audit (Algemene Rekenkamer, 2022) has also been complaining for years that there is hardly any information available to help it perform its work properly, that is, to verify whether the central government spends public money sensibly, economically, and carefully. In this regard, it accuses the minister of laxity and vagueness. It has further criticism as well. For example, she points to the lack of a well-founded theory regarding the origins of educational disadvantages and of a coherent vision for national policy; according to her, there is an arbitrary collection of topics, coupled with the absence of more concrete objectives (Tweede Kamer, 2000).

According to the Ministry of Finance (IBO, 2017), primary schools primarily use their EDP funds for class size reduction and teaching assistants, followed by a second teacher per class. However, these are not the interventions that score highest on cost-effectiveness. Schools invest less in interventions such as pre-teaching (providing instruction to weak students in advance), parental involvement, and remedial teaching, even though these interventions perform relatively favorably in research.

The National Education Program (NEP) offers a recent glimpse into the spending of the weighted funds. This program was hastily set up in 2021 by the minister to mitigate the expected negative effects of school closures during the

COVID-19 pandemic (Driessen, 2026). In total, the unprecedented sum of 8.5 billion euros was reserved, of which 5.8 billion was for primary and secondary education. The allocation for the latter sectors was based on the criteria of the weighted funding scheme. Schools received hundreds of thousands to millions of euros that they had to spend on evidence-based interventions, that is to say, interventions whose effectiveness has been established through rigorous scientific research. Schools with many weighted students received up to nearly double the amount in additional subsidies. The ministry had drawn up a list of 24 approved interventions, the so-called Menu Card, from which schools were required to choose. These subsidies had to be spent within two years; if this was not achieved, they had to be repaid. The Court of Audit (Algemene Rekenkamer, 2022), pointed, however, to the fact that this did not fall under the minister's competences. The extra subsidies were paid out as part of the lump sum, which implies that schools are free to spend them as they see fit and that no time limit can be imposed.

Because knowledge regarding proven interventions was lacking within the ministry, the minister resorted to the English EEF (Education Endowment Foundation). That organization possesses—by its own account—an overview of proven interventions, from which the ministry, supported by a group of experts, made a selection. However, further research revealed that these interventions were not nearly as proven as suggested (Driessen, 2026). The evidence was often flimsy, all cited studies were lumped together, and an average 'effect' was calculated. However, based on that unclear effect (Effect Size), schools were forced to make a choice. The main thing the underlying studies had in common was a vague, general designation, such as 'feedback' or 'sport'. However, these studies differed enormously in terms of the concrete implementation of the intervention, the specific target group, and the age category. The design and quality of the studies themselves also varied widely. The result was an average effect with an unacceptably large variation and margin of uncertainty. What the minister evidently forgot is that it is not only a question of what works, but also for whom, and under what conditions (Dues, 2020). He therefore obliged schools to invest billions in interventions for which there was little or no evidence that they were effective. It should be mentioned that in a relatively short time the terminology changed, from 'evidence-based', to 'evidence-informed', and then 'promising'.

According to the Court of Audit (Algemene Rekenkamer, 2021), all primary schools they interviewed did not use the Menu Card as a tool to select interventions, but primarily as an accountability instrument to the ministry. School leaders did not arrive at different choices, but attempted to justify the choices already made (in retrospect) as interventions from the Menu Card. Schools indicated that they mainly wanted to spend the subsidy on interventions that, according to the menu, are relatively expensive but not truly effective, such as hiring teaching assistants and reducing class sizes.

Questionnaire research by Van den Broek et al. (2025) reveals what, according to primary and secondary school leaders, was primarily done with the NEP funds. This is: instruction in small groups (primary education: 85%; secondary education: 79%); deployment of teaching assistants (79%; 76%); interventions aimed at well-being (70%; 90%); direct instruction (48%; 0%); 1-on-1 guidance (47%; 65%); sports activities (0%; 75%). It is noteworthy that two of the most chosen interventions call for additional teachers and supervision. Even more than usual, there was a dire shortage of teaching staff during the NEP period, particularly at disadvantaged

schools. This was exacerbated by the fact that new staff moved primarily to non-disadvantaged schools, making the situation at disadvantaged schools even more critical. It is also striking that no data were available regarding the effectiveness of 'well-being' and 'direct instruction', which raises the question of the extent to which school leaders actually considered the criterion of 'proven effective' during their selection. In general, the average effect of the interventions was not particularly strong (apart from the criticisms described above).

3.2. Mother tongue education

An important policy track of the EDP concerned Mother Tongue (and Culture) Education (MTE), which was introduced in 1974 (Driessen, 2005). It was originally intended for guest worker children and set up based on the concept of return: guest workers were here temporarily and would eventually return to their country of origin. To facilitate their children's easy re-entry into education there, they could already attend subsidized lessons in their own language and culture for up to 5 hours per week. The largest target groups were Turkish and Moroccan children, but OETC was also offered to numerous smaller language groups, such as Spanish, Greek, and Italian. When it became apparent that the guest workers were not remigrating but were staying in the Netherlands, the policy objectives changed. The focus shifted to reducing educational disadvantage. The underlying idea was that children would first have to learn their mother tongue in order to successfully acquire Dutch. Over the years, the objectives and design changed again. First, the cultural component was discontinued, and subsequently, it increasingly functioned as a support language in Dutch education; additionally, the aim was to teach the mother tongue and thereby maintain contact with one's own culture. In 2004, the subsidy for this education was discontinued due to a lack of evidence of positive effects on integration and language development.

It is unclear how much money was spent on MTE during its thirty years of existence. A rough estimate puts it at between 1.5 and 2 billion euros (cf. Algemene Rekenkamer, 2001). From the outset, MTE was plagued by problematic circumstances, starting with the constantly changing objectives, which were also particularly vague. Importantly, policy and practice showed little agreement. This is largely due to the lack of knowledge within the department. MTE had to be taught in the official national language. This created an extremely peculiar situation for a very large proportion of the children. Thus, the Moroccan Berber-speaking children were required to learn official Standard Arabic, a foreign language to them. Something similar applied to, for example, the Kurdish children who had to learn Turkish, and children from the Catalonia region who had to learn Spanish. It was downright baffling that the responsible Minister of Education only discovered later that the Berber children did not speak Arabic at all, but had to resort to Dutch as a lingua franca to communicate with one another (Oosterbaan, 1992). Apparently, his assistants at the department did not know this either, or at least they had not told him. This lack of knowledge within the department meant that perhaps as many as half of these target group children had to learn a foreign language, and that for them, therefore, the thesis (or axiom) that one must first master one's mother tongue in order to learn Dutch well could not, by definition, hold true. The majority of this so-called mother tongue education was provided during regular school hours, which means that all those hours dedicated to learning a foreign language were deducted from other subjects. While everyone had been aware of this malpractice for many years, no one intervened and the ministry continued to pay out substantial subsidies. Apart from all

this, there were various practical implementation problems from the very beginning. For instance, many MTE teachers spoke little or no Dutch, often lacked a (Dutch) teaching qualification, there was no integration into the Dutch team, there was excessive absenteeism due to illness, and there were no adequate teaching materials and classrooms.

3.3. Early childhood education and care

Projects aimed at preventing educational disadvantages among preschoolers were already being implemented in the Netherlands in the 1970s. These took place in a few major cities. Around 2000, however, a policy was initiated at the national level: ECEC, for children aged 2.5 to 6 growing up in socio-economically disadvantaged situations. Funding for ECEC by the ministry is based on the criteria of the weighting scheme. However, it is downright confusing and inconsistent that municipalities, which are responsible for the implementation of the preschool phase, are allowed to determine their own target group criteria. Consequently, the children for whom national subsidies are granted are by no means always the same children who constitute the target groups within municipalities. Alongside the weighting scheme, ECEC is currently the most important instrument the government uses to prevent and combat educational disadvantages.

Attention to ECEC gained momentum around 2000 in response to the recommendations of the Meijnen Commission. It was advised to intervene early, under the motto: the sooner you start, the more progress you can make, and to develop targeted programs accordingly. That advice was based on a study by Leseman and Cordus (1994), which referred to the work of Nobel laureate James Heckman. Building on this, various early childhood education programs have been developed by commercial institutions, most for implementation in daycare centers, playgroups, and kindergarten groups, some (also) for parents at home. Since then, Heckman has been cited almost as standard (and uncritically, even by the ministry) as proof of the effectiveness of such programs and thereby justification for the annual subsidization of hundreds of millions of euros.

Recently, I have examined Heckman's work closely, and in particular the (model) programs he analyzed that form the basis for his unbridled optimism (Driessen, 2025). That turned out to have an extremely narrow and shaky foundation. The programs examined, particularly Perry and Abecedarian, and the conditions under which they were implemented are in no way comparable to the Dutch ECEC situation. Both programs were implemented over half a century ago under completely different circumstances, with entirely different target groups, tiny sample sizes, full-time intensity, better-qualified supervision, and a completely unrealistic price tag (Perry cost $43,500 per child, and Abecedarian no less than $120,000; the Dutch preschool phase costs a maximum of $10,000 per child).

The programs developed here are assessed on their merits by the Netherlands Youth Institute (NJi), and then assigned or withheld the label ‘recognized.’ There are currently 9 of them (NJi, 2026). The NJi Recognition Committee discerns four levels of recognition: (1) theoretically well-founded; (2) initial indication of effectiveness; (3) good indication of effectiveness; (4) strong indication of effectiveness. For levels 1–3, only 2 studies with strong to very strong evidence are required, of which at least one must be a Dutch study. Of the 9 recognized programs, only one has achieved level 2; all other programs are still at level 1. The ministry advises ECEC institutions to use a recognized program, but cannot make it mandatory. Institutions may also use other programs, but these programs have not been assessed and would likely not even reach

level 1. In other words: the ministry has spent many billions on ECEC over the past 25 years, without there being any real evidence that a single program is effective.

4. What has it all yielded

Another particularly difficult question. Perhaps incomprehensible, but it must actually be concluded that nothing can be said about the effectiveness of the EDP. Normally, one would expect that new policy is based on a thorough literature review, that practical implementation and feasibility have been examined in detail, that concrete objectives have been formulated, that a (small-scale) pilot has been conducted where possible, and that it has been established in advance how it will be investigated to what extent the objectives will ultimately be achieved. I have already indicated above that there have been frequent complaints about vague, broadly interpretable, and unspecific objectives, which were also modified over the years and were invariably accompanied by unclear and often contradictory communication (Algemene Rekenkamer, 2021; Meijnen, 2004; Mulder, 1996). Although the Education Priority Policy and the NPO have been accompanied by their own research programs from the outset, these were designed in such a way that it was not possible to make causal statements about their effectiveness. As a result, the studies conducted within the framework of the EDP were descriptive or correlational, and their effectiveness could not be established. Nevertheless, I will attempt below to examine what this has yielded, or perhaps more correctly, what the current state of affairs is.

4.1. The weighting system

Mulder's dissertation (1996) was devoted to the evaluation of the EDP for the period 1988-1992. She compared the performance of native and non-native target group children with that of non-target group children. Her conclusion was that seven years after the start of the Education Priority Policy, non-native target group students (particularly Turkish and Moroccan) still lagged behind native target group children, who in turn lagged behind non-target group students. The cautious final assessment was that the extra resources had barely led to improved learning performance; the structural disadvantage of these students remained significant.

In an update, several years later, Meijnen (2004) concluded regarding the period 1997-2002 that the position of non-native target group children in primary education appeared to be developing positively, although Turkish and Moroccan students still lagged far behind. Although native target group students performed somewhat less unfavorably, little progress could be observed among them; rather, a slight decline. No connection could be established with the policy. In response to these results, Meijnen pointed out, once again, that policy changes were not supported by theoretical and empirical arguments and that the policy still lacked a clear means-ends analysis regarding the realism of the objectives.

A few years later, Driessen (2012) mapped the performance developments in primary education. In the period 1995-2011, a substantial reduction in delays took place. This varied roughly from an improvement of 25% for language proficiency to 15% for numeracy skills. However, the progress of the native target group was considerably smaller than that of the non-native target group. The native target group had even declined in terms of language and math. The immigrant target group had actually improved significantly on those two skills (40% and 32%, respectively). In the period 2003-2011, the total target group made 15% progress. However, here too, that progress had to be largely attributed to the immigrant target group. For the period 2008-2011, the picture was considerably less favorable. For the total target

group, there was even a decline in most respects. From the breakdown by native and immigrant population, it became clear that this decline was primarily caused by the immigrant target group. Moreover, these negative developments do not apply only to the Netherlands, but are widely observed (Jorde Sandsør, 2023; Strand, 2025; Wexler, 2019).

Recently, Haelermans and Havermans (2026) conducted a number of quantitative analyses on regional, national, and international datasets as part of the NEP evaluation. This involved both cognitive (reading, spelling, and math) and non-cognitive (creativity, curiosity, critical thinking, and problem-solving) measures. Importantly, the focus was on developments over time (pre-COVID-19, COVID-19, and post-COVID-19), but no information was available regarding NEP interventions. It is therefore not possible to establish a relationship between the two, let alone a causal one. Although there was considerable variation between school years, sectors, grade levels, and subjects, Haelermans and Havermans conclude that students are slowly catching up on their COVID-related learning deficits, although this applies to a greater extent to primary than to secondary education. Furthermore, there are gender and socio-economic background differences: boys perform better than girls, and students with highly educated parents perform better than students with moderately or less educated parents. For secondary education, the situation differed from that of primary education. There, the decline in performance continued after the pandemic compared to during the pandemic. Cognitive skills declined after and during the pandemic compared to before the pandemic, particularly in primary education. In secondary education, students score higher on some non-cognitive measures, while scoring lower on others. In summary, the researchers conclude that, in many respects, students have not yet reached the pre-pandemic level after the pandemic. It is downright disappointing that no relationship whatsoever can be established between the NEP funds awarded to schools, the interventions implemented, and the outcomes. We will never know whether the huge subsidies disbursed under the weighting scheme yielded anything. Nor will we know whether the so-called proven interventions had any effect.

4.2. Mother tongue education

The number of studies into (the effectiveness of) MTE is extremely limited; it mainly concerns a few descriptive and correlational studies. Given what I have written above regarding the conditions of that education, it should be clear that little could be expected from it from the outset. The vagueness and contradiction of the goals certainly contributed to this. A large-scale study concluded that the Spanish children learned the most about their language and culture, the Turkish children did reasonably well, and the Moroccan children learned nothing from the lessons in their 'own' language (which is logical, since the language to be learned was actually a foreign language). Other studies indicated that children who attended mother tongue lessons generally did not perform worse in terms of Dutch language than children who did not. However, positive 'effects' on Dutch were not found (Driessen, 2005). This implies that mother tongue education has also not contributed to resolving educational disadvantages.

4.3. Early childhood education and care

It is completely unclear what the contribution of ECEC is to preventing and combating educational disadvantages. Reference is often made to foreign studies, particularly from the US, but it has been shown that for many of them, the conditions

are incomparable to those of Dutch ECEC, or that the programs themselves differ significantly, the evaluations are methodologically weak, and the 'effects' are not strong, spread out in all directions, and fade away over time (Driessen, 2025; Lu, 2025). A meta-analysis of all Dutch studies (Fukkink et al., 2017) showed that the average effect of these studies was 0 (zero). It should be noted, however, that none of the underlying studies met strict methodological requirements. Thus, for ECEC as well, there is no evidence that it is effective.

5. And now?

5.1. Conclusions

What can be learned from the above? Since 1974, government policy has been pursued to combat educational disadvantages resulting from the socio-economic circumstances in which children grow up. However, that policy is inconsistent, changes constantly (primarily due to party political views), lacks a solid theoretical and empirical foundation, and is devoid of the necessary knowledge among government ministers and policymakers. There also appears to be an unwillingness among ministers to accompany policy with adequate research in order to adjust it—if necessary—based on that research. In addition, there is a significant mismatch between policy and practice. The latter is primarily caused by the persistent vagueness of policy and the lack of concrete objectives, and, of course, the far-reaching freedom that school boards enjoy in implementing national policy.

Unfortunately, all the tens of billions in subsidies have yielded nothing (demonstrable)—on the contrary. Despite all those huge sums, educational performance seems to have been declining for years. The allocation and expenditure are based on shaky ground. And the results reflect that. Therefore, it seems like it is time to give this careful thought and come up with a suitable alternative. It is sometimes argued: yes, but you do not know what would have happened without all that money (cf. Coe, 2017). Indeed, but you know, at the very least, that the approaches devised and implemented so far have offered insufficient relief. Simply continuing on the old path, or even throwing more subsidies at it, does not seem like the most obvious solution.

What stands out is that in those more than half a century, lamentably little adequate research has been conducted; there is a distinct lack of impact evaluations, that is to say: verifying through thorough research whether the objectives pursued by the policy have actually been achieved. The unimaginable consequence is that the government pours tens of billions of euros into policy, yet has no idea whether it has been effective. The student performance level, which has been declining significantly for quite some time, gives cause for the worst fears. All that money down the drain? A question that inevitably arises, or rather ought to arise, is why the ministry is not commissioning a proper investigation. Is it not interested? Does it simply not want to know whether the policy has been effective?

What is truly incomprehensible is that the ministers knew, or ought to have known, that their assumptions were flawed, and yet continued with their policy, without (actively) informing the field of this (nor the House of Representatives). We saw this at the MTE (half of the children had to learn a foreign language instead of their mother tongue; the transfer from 'mother tongue' to the Dutch language was therefore a pipe dream and, in any case, completely unproven), at the NPE (the ministry could not oblige schools to choose from its list of interventions, nor to use up the subsidies within two years; and the interventions were not nearly as proven as

suggested), and at ECEC (the scientific basis was lacking; proof that VVE works in the Netherlands is completely absent). These are all cases where the ministry has not done its homework properly, where there is therefore an evident lack of knowledge and/or a critical attitude, or where the ministers simply pushed through their will, in defiance of the facts, or precisely because of the lack thereof (cf. Onderwijsraad, 2026). The ministers stoically clung to the path once chosen. This affects – to begin with – hundreds of thousands of children, year after year. And I am not even talking about all that money that could have been better spent.

But undeniably, the blame also lies with science. Why does hardly anyone report that the weighting system is far too crude and that it is questionable where the money ends up and how it is spent? Why did (linguistic) scientists not sound the alarm and announce to the parties involved that many non-native speaking children were not required to learn their mother tongue via the OETC, but a foreign language? Why has no one ever taken the trouble to delve into the writings of James Heckman and examine what we can really use them for here and now? Why did the group of experts that oversaw the introduction and implementation of the NPE not first conduct thorough research into the mandatory evidence-based interventions, which were easily demonstrably not proven at all? Is there a lack of methodological knowledge? Where is the objective and critical-scientific perspective on the entire state of affairs? Is this an ideologically colored battle of directions? Is it taboo to be critical (certainly not towards potential clients for research)? Is it simply a matter of preserving jobs? (Also see De Bruyckere, 2026.)

5.2. From group to individual?

All in all, the analyses show that there are serious grounds to doubt the current substantiation of the most important instrument of the EDP, namely the weighting scheme, and the distribution of additional financial resources based on it. The question therefore is whether the current group-based approach to the problem of disadvantage is (still) justified and whether serious consideration should not be given to alternatives. Until now, the individual-based approach involving assessment has been rejected, particularly because it was assumed that individual assessment would be expensive and would lead to unreliable results in very young children. However, it is highly questionable whether this approach, possibly with a few adjustments, is not actually better than the current weighting scheme (cf. Leseman, 2004). If the problem were the leapfrog development of preschoolers, this would argue for more frequent assessment in order to gain a better grasp of their development. According to some linguists, suitable language tests for young children are indeed available (Onderwijsraad, 2002; Van den Branden, 2007; Verhoeven & Vermeer, 2005). However, other options can also be considered. In early childhood education, extensive preschooler monitoring systems have been used intensively for years (often utilizing a combination of standardized tests and observations by daycare workers and teachers). A two-stage approach could also be considered, in which an initial screening by daycare workers and teachers is followed by more extensive testing, or an initial selection based on family characteristics followed by testing (cf. Onderwijsraad, 2001).

Regarding language tests, I would like to refer to the comprehensive review study by Colpin et al. (2006), which clearly outlines the ins and outs. They also conclude that testing toddlers and preschoolers aged 6 and under is not without problems. But when this ideal is not feasible, why not choose a next-best option, that is, administering a test at age 6, at the moment the children enter grade 3? Such

a test would have to meet several specific conditions. In addition to good reliability, it should have high predictive validity, meaning it should accurately predict later performance. Furthermore, it should differentiate well between the target group and the non-target group. More specifically, it should distinguish between children who underperform due to unfavorable circumstances in the environment in which they grow up and children whose low performance is the result of mental, physical, or behavioral factors inherent in the child itself. Both groups require their own form of extra attention.

Family characteristics (as operationalized by the Ministry of Education) do not form a good basis for the allocation of extra subsidies. One could also go further and do nothing at all with that background information, looking only at actual functioning, or actual performance, respectively. Van Schooten et al. (2007) assessed a number of tests for young children on their merits and showed that while some (sub)tests performed well in most respects, they did not—when strict standards were applied—regarding bias (distinguishing between different groups). However, with a somewhat more lenient approach, only the age difference remained in this regard: there was no problem with children older than 6, but there was with children younger than 6.

One of the tests Van Schooten et al. examined was the OBIS (Primary Grades Information System; Van der Hoeven-Van Doornum, 2005), an adaptive, computer-controlled screening tool and tracking system for grades 1 and 2 (ages 4 and 5), which, in addition to passive vocabulary, sound articulation and sound discrimination, has thirteen other subtests, six of which are intended as language tests and seven as arithmetic tests. The OBIS also provides detailed diagnostic information at subtest level. In a psychometric sense, the test did well; in addition, test taking did not lead to any problems in practice. The administration time is approximately 20 minutes. The material costs at the time – when testing 50 children per school – amounted to €400; No personnel input is required to process the data.

This OBIS test is a Dutch adaptation of the English PIPS (Performance Indicators in Primary Schools; Tymms et al., 2014), which maps the development of children from their start at primary school to the transition to secondary education (a period of 6 years). The PIPS has very high internal consistency ($\alpha = 0.92 - 0.94$), test-retest reliability ($r = 0.98$), and predictive validity ($r = 0.70$ or 50% explained variance). The PIPS tests have been administered more than five million times worldwide (including England, Scotland, New Zealand, Australia, and South Africa). In the meantime, the PIPS has been further developed and transitioned into the BASE (Reception Baseline

Assessment; Cambridge Insight, 2026; Jellis, 2022), a test for 4- and 5-year-olds that is administered almost half a million times in 125 countries every year. This test includes the following components: Literacy, Numeracy, Communication & Language, and Personal Social and Emotional Development. A question that immediately arises is of course: are these tests able to generate a reliable state of affairs for such young children? Considering the millions of purchases, the education world seems content. Therefore, it would be a good idea to take a closer look at this and examine what these types of tests could mean for educational disadvantage policy, not only for well-founded financing thereof, but also for educational practice itself.

A relevant question in all of this is also how long additional funding per student should last. You could say that a child growing up in unfavorable circumstances cannot be expected to improve quickly, and that additional funding is therefore

obvious throughout the school period, even when performance no longer deviates from that of the non-target group. On the other hand, there is also something to be said about the fact that, when a student no longer shows performance deficits, he or she no longer needs additional financial support, or at least not until tests show that the child has fallen behind again or is in danger of falling behind again. This assumes that regular (formative) testing continues, for example via a student monitoring system. Additional financing could also be linked to a specific period, for example, up to two years after arrears were last identified.

In a study into ECEC (Driessen et al., 2015), practical experts suggest completely abandoning the way in which this is currently structured. Instead, they advocate an Inclusive ECEC approach (analogous to Inclusive Education), an integrated provision that is accessible to all children. The approach is a development-oriented educational policy, with a differentiated offering tailored to the specific capabilities and needs of each child. All factors that in any way hinder the development of the child are important. This not only concerns children with 'traditional' disadvantages and care needs (the target groups of the Educational Disadvantage Policy and Special Education Needs, respectively), but also children who have more to offer (Excellence Policy). Such an individual-oriented approach ('inclusive education') could possibly also be applied in primary education. Compared to the way in which the Educational Disadvantage Policy is currently being shaped, such an integrated approach would represent a radical break.

The above is certainly not a ready-made solution. It concerns some suggestions for possible alternatives to the current weighting system, which still require a lot of further research. So be it, anything is better than continuing to struggle in the current way.

Author contributions: Not Applicable

Funding: Not Applicable

Ethical approval: Not applicable.

Informed consent statement: Not applicable.

Conflict of interest: Not applicable.

References

1. Algemene Rekenkamer. Onderwijs in Allochtone Levende Talen (Oalt). Algemene Rekenkamer; 2001. https://www.rekenkamer.nl/site/binaries/site-content/collections/documents/2001/04/10/onderwijs-allochtonen-levende-talen/Onderwijs+Allochtonen+Levende+Talen.pdf
2. Algemene Rekenkamer. Aandachtspunten bij het Nationaal Programma Onderwijs. Algemene Rekenkamer; 2021. https://www.rekenkamer.nl/documenten/2021/03/31/aandachtspunten-bij-het-nationaal-programma-onderwijs
3. Algemene Rekenkamer. Verantwoordingsonderzoek. Ministerie van Onderwijs, Cultuur en Wetenschap (VIII). Rapport bij het Jaarverslag 2021. Algemene Rekenkamer; 2022. https://www.rekenkamer.nl/documenten/2022/05/18/resultaten-verantwoordingsonderzoek-2021-ministerie-van-onderwijs-cultuur-en-wetenschap
4. Cambridge Insight. Reception Baseline Assessment (BASE) for ages 4-5. Pre-school, Kindergarten, Reception baseline assessment. Available online: https://www.cambridge.org/insight/reception-baseline-assessment (accessed on 10 August 2026).
5. CBS. Onderzoek herijking risico-indicator onderwijsachterstanden basisonderwijs. Fase 2. CBS; 2026. https://dashboards.cbs.nl/longreads/2026/onderzoek-herijking-risico-indicator-onderwijsachterstanden-basisonderwijs-fase-2/
6. Coe R. Evaluation: Why, what and how. Blog Impact, 20 May 2017. Available online: https://my.chartered.college/impact_article/evaluation-why-what-and-how/ (accessed on 12 August 2026).

7. Colpin M, Gysen S, Jaspaert K, et al. Studie naar de wenselijkheid en haalbaarheid van de invoering van centrale taaltoetsen in Vlaanderen in functie van gelijke onderwijskansen. K.U. Leuven; 2006. https://www.arts.kuleuven.be/cto/onderzoek/onderzoek-evaluatie/wenselijkheid-centrale-taaltoetsen/colpin-ea-2006-haalbaarheidsstudie-centrale-taalto.pdf
8. De Bruyckere P. Defending science means talking about its problems. Available online: https://theeconomyofmeaning.com/2026/09/07/defending-science-means-talking-about-its-problems/ (accessed 7 September 2026).
9. Driessen G. From cure to curse: The rise and fall of bilingual education programs in the Netherlands. In J. Söhn (Ed.), The effectiveness of bilingual school programs for immigrant children. WZB Discussion Paper SP IV 2005-601 (pp. 77-107). Wissenschaftszentrum Berlin für Sozialforschung; 2005. https://doi.org/10.5281/zenodo.10825262
10. Driessen G. De doelgroepen van het onderwijsachterstandenbeleid: Ontwikkelingen in prestaties en het advies voortgezet onderwijs. ITS; 2012. https://doi.org/10.5281/zenodo.10825615
11. Driessen G. De wankele empirische basis van het onderwijsachterstandenbeleid. De afnemende validiteit van indicatoren voor de toewijzing van extra middelen. Mens & Maatschappij. 2015; 90(3): 221-243. https://doi.org/10.5117/mem2015.3.drie
12. Driessen G. The validity of educational disadvantage policy indicators. Educational Policy Analysis and Strategic Research. 2017; 12(2): 93-108. https://doi.org/10.5281/zenodo.6803502
13. Driessen G. The many facets of educational disadvantage. Policies, interventions, effects. Eliva Press; 2022. https://doi.org/10.5281/zenodo.6641161
14. Driessen G. The fragile foundation of pre- and early-school programs for disadvantaged children. Forum for Education Studies. 2025; 3(1): 1869. https://doi.org/10.59400/fes1869
15. Driessen G. Evidence-based educational interventions as the solution to learning gaps caused by the COVID-19 pandemic? Research on Preschool and Primary Education. 2026; 4(1): 57-74. https://doi.org/10.55976/dma.42026152657-74
16. Driessen G, Veen A, Van Daalen M. VVE-doelgroepkinderen in de voorschoolse fase. Indicering en aanbod. ITS; 2015. https://doi.org/10.13140/2.1.1207.8883
17. Dues J. Why evidence-based practices don't work: Part II. Thomas B. Fordham Institute; 2020. https://fordhaminstitute.org/national/commentary/why-evidence-based-practices-dont-work-part-ii
18. Fukkink R, Jilink L, Oostdam R. A meta-analysis of the impact of early childhood interventions on the development of children in the Netherlands: An inconvenient truth? European Early Childhood Education Research Journal. 2017; 25(5): 656-666. https://doi.org/10.1080/1350293X.2017.1356579
19. Geurtsen K. Toetscultuur zet jonge kinderen onder druk: 60 procent leraren schat in dat leerlingen buiten school trainen. Available online: https://pointer.kro-ncrv.nl/toetscultuur-zet-jonge-kinderen-onder-druk-60-procent-leraren-schat-in-dat-leerlingen-buiten-school (accessed on 16 August 2026).
20. Goorhuis S. Peuters en kleuters zijn niet toetsbaar. Available online: https://wij-leren.nl/kleuters-toetsen.php (accessed on 18 August 2026).
21. Haelermans C, Havermans W. Academic resilience and policy reform in the Netherlands. In L Volante K De Witte, O Giancola et al. (Eds.), Risk, resilience, and recovery across global education systems: The long-term impact of pandemic reforms (pp. 110-128). Routledge; 2026. https://doi.org/10.4324/9781003608806-9
22. Hunt E, Perrott M, Jiménez E, et al. Growing apart. The evolution of the disadvantage gap. EPI/EEF; 2026. https://epi.org.uk/publications-and-research/growing-apart-the-evolution-of-the-disadvantage-gap/
23. IBO. Onderwijsachterstandenbeleid, een duwtje in de rug? Interdepartementaal beleidsonderzoek naar het onderwijsachterstandenbeleid. Ministerie van Financiën; 2017. https://archief.rijksbegroting.nl/system/files/12/ibo-onderwijsachterstandenbeleid-eindrapport-een-duwtje-de-rug.pdf
24. Jellis C. Progress in the first year at school. Research Matters. 2022; 34: 42–56. https://www.cambridgeassessment.org.uk/Images/research-matters34-progress-in-the-first-year-at-school.pdf
25. Jorde Sandsør A, Daae Zachrisson H, Karoly L, et al. The widening achievement gap between rich and poor in a Nordic country. Educational Researcher. 2023; 52(4): 195-205. https://doi.org/10.3102/0013189X221142596
26. Kalmijn M, Kraaykamp G. De verklaarde variantie verklaard: Een vergelijking van sociologische onderzoeksartikelen in de periode 1975-1998. Mens & Maatschappij. 1999; 74(2): 166-179. https://matthijskalmijn.nl/onewebmedia/Kalmijn%20Kraaykamp%20-%20MM%20-%20Verklaarde%20variantie.pdf
27. Kloprogge J. Comment. LinkedIn. Available online: https://www.linkedin.com/feed/update/urn:li:activity:7474402621286551552/?dashCommentUrn=urn%3Ali%3Afsd_comment%3A%28747442171087681536 0%2Curn%3Ali%3Aactivity%

3A7474402621286551552%29 (Accessed on 21 June 2026).
28. Ladd H, Fiske E. Weighted student funding in the Netherlands: A model for the U.S.? Journal of Policy Analysis and Management. 2011; 30(3): 470-498.
29. https://doi.org/10.1002/pam.20589
30. Leseman P. De toegevoegde waarde van vroeg testen. Pedagogiek. 2004; 24(1): 3-11. https://journal-archive.aup.nl/pedagogiek/vol_24_nr_1_-_forum_-_de_toegevoegde_waarde_van_vroeg_testen.pdf
31. Leseman P, Cordus J. (Allochtone) kleuters meer aandacht: Advies van de commissie (Voor)schoolse educatie in opdracht van de Minister van Welzijn Volksgezondheid en Cultuur en de Minister en Onderwijs en Wetenschappen. Ministerie van Welzijn, Volksgezondheid en Cultuur/ Ministerie van Onderwijs en Wetenschappen; 1994.
32. Liu J, Peng P, Zhao B, et al. Socioeconomic status and academic achievement in primary and secondary education: A meta-analytic review. Educational Psychology Review. 2022; 34: 2867–2896. https://doi.org/10.1007/s10648-022-09689-y
33. Lu M. Evaluating Head Start Program for low-SES students: Effectiveness, challenges, and future directions. Lecture Notes in Education Psychology and Public Media. 2025; 114(1): 8-13. https://www.researchgate.net/journal/Lecture-Notes-in-Education-Psychology-and-Public-Media-2753-7056?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InB1YmxpY2F0aW9uIiwicGFnZSI6InB1YmxpY2F0aW9uIn19
34. Meijnen W. Onderwijsachterstanden. Update van een review. Recente ontwikkelingen inzake onderwijsachterstanden. SCO-Kohnstamm Instituut; 2004. https://www.nko.nl/sites/nro/files/migrate/Onderwijsachterstanden-update-van-een-review.pdf
35. Mulder L. Meer voorrang, minder achterstand? Het Onderwijsvoorrangsbeleid getoetst. ITS; 1996.
36. https://repository.ubn.ru.nl/bitstream/handle/2066/272925/mmubn000001_219267502.pdf
37. Nji. VVE-programma's. Available online: https://www.nji.nl/zoeken?keyword=&type=intervention&age_upper_limit=2&age_lower_limit=6&subject%5B39%5D=39 (accessed on 20 August 2026).
38. Onderwijsraad. Voorschools en buitenschools. Onderwijsraad; 1998. https://www.raadrvs.nl/site/binaries/site-content/collections/documents/1998/06/01/voorschools-en-buitenschools/Voorschools_en_buitenschools.pdf
39. Onderwijsraad. Wat 't zwaarst weegt... Een nieuwe aanpak voor het Onderwijsachterstandenbeleid. Advies. Onderwijsraad; 2001.
40. Onderwijsraad. Over leerlinggewichten en schoolgewichten. Advies. Onderwijsraad; 2002. https://onderwijsraad.sitearchief.nl/?timestamp=20190923054357&url=http%3A%2F%2Fonderwijsraad.esn.nl%2Fpublicaties%2Farchief%2Fitem21&dl=t#archive
41. Onderwijsraad. Onderzoekinfrastructuur voor onderwijsbeleid. Onderwijsraad; 2026. https://www.onderwijsraad.nl/documenten/2026/07/09/onderzoeksinfrastructuur-voor-onderwijsbeleid
42. Oosterbaan W. Twintig jaar onderwijs in eigen taal en cultuur. Kluwen van goede bedoelingen. NRC, 22 October 1992. Available online: https://www.nrc.nl/nieuws/1992/10/22/twintig-jaar-onderwijs-in-eigen-taal-en-cultuur-kluwen-7160368-a631251 (accessed 31 Augsut 2026).
43. Posthumus H, Bakker B, Van der Laan J, et al. Herziening gewichtenregeling primair basisonderwijs – Fase I. CBS; 2016. https://www.cbs.nl/nl-nl/publicatie/2017/06/herziening-gewichtenregeling-primair-onderwijs
44. Sirin S. Socioeconomic status and academic achievement: A meta-analytic review of research. Review of Educational Research. 2005; 75(3): 417-453.
45. https://doi.org/10.3102/00346543075003417
46. Smedslund G, Arnulf J, Smedslund J. Is psychological science progressing? Explained variance in PsycINFO articles during the period 1956 to 2022. Frontiers in Psychology, 2022; 13:1089089. https://doi.org/10.3389/fpsyg.2022.1089089
47. Strand S. Written evidence submitted to the Independent Inquiry into White Working Class Educational Outcomes. Available online: https://www.researchgate.net/publication/405856091_Written_evidence_submitted_to_the_Independent_Inquiry_into_White_Working_Class_Educational_Outcomes (accessed on 2 September 2026).
48. Tweede Kamer. Aanpak onderwijsachterstanden. KMST 51084. Sdu Uitgevers; 2000. https://www.rekenkamer.nl/site/binaries/site-content/collections/documents/2001/01/30/bestrijding-van-onderwijsachterstanden/Bestrijding+van+onderwijsachterstanden.pdf
49. Tymms P, Merrell C, Hawker D, et al. Performance Indicators in primary schools: A comparison of performance on entry to school and the progress made in the first year in England and four other jurisdictions. Research report. Department for Education; 2014. https://www.gov.uk/government/publications/performance-indicators-in-primary-schools

50. Van den Branden K. Taaltoetsen en schoolbeleid. Is meten echt beter dan weten? Vonk. 2007; 37(1): 3-20. https://www.arts.kuleuven.be/cto/onderzoek/onderzoek-evaluatie/wenselijkheid-centrale-taaltoetsen/vonk-jrg37-002-1-taaltoetsen-en-schooltaalbeleid.pdf
51. Van den Broek A., Termorshuizen T, Luyten E, et al. Evaluatie Nationaal Programma Onderwijs – funderend onderwijs. ResearchNed; 2025. https://www.researchned.nl/publicatie/ evaluatie-nationaal-programma-onderwijs-funderend-onderwijs/
52. Van den Burg-Poortvliet N. Kleuters toetsen mag straks niet meer: wat dan wel? Beter Begeleiden. 2021; 11(2): 1-4. https://www.lbbo.nl/wp-content/uploads/2021/03/KleuterartikelDEFLR.pdf
53. Van der Hoeven-Van Doornum A. Development on scale, instruction at measure. OBIS, a system of value-added indicators in primary education. ITS; 2005.
54. Van Schooten E, Smeets E, Driessen G. Taaltoetsen voor taalzwakke leerlingen in het primair onderwijs. SCO-Kohnstamm Instituut/ITS; 2007. https://www.vosabb.nl/wp-content/uploads/2013/06/Bijlage_taaltoets.pdf
55. Verhoeven L, Vermeer A. Het ongelijk van Netelenbos? Toetsing van kleuters en hun prestaties op de Cito Eindtoets Basisonderwijs. Toegepaste Taalwetenschap in Artikelen, 2005; 74(1): 123-133. https://doi.org/10.1075/ttwia.74.12ver
56. Wexler N. The achievement gap hasn't budged In 50 years. Now what? Forbes. 18-03-2019. https://www.forbes.com/sites/nataliewexler/2019/03/17/the-achievement-gap-hasnt-budged-in-50-years-now-what/#4cf2b84e4d90
57. Zhao B, Liu J, Luo L. Starting well: A meta-analysis of the relationship between socioeconomic status and academic skills in early childhood. Review of Educational Research. 2025; 95(6): 1414-1449. https://doi.org/10.3102/00346543241293448