



De laatste jaren is er terecht veel aandacht voor de doorstroomtoets en de cruciale rol die deze speelt in de selectie en advisering van leerlingen naar het voortgezet onderwijs. Een minstens even belangrijk aspect is de andere, sturende functie van de toets: het bepalen van de onderwijsresultaten van een school in het zogenoemde onderwijsresultatenmodel van de Inspectie van het Onderwijs. Een pleidooi voor herziening van deze systematiek.

DE OPEENSTAPELING VAN PROBLEMEN MET HET ONDERWIJSRESULTATENMODEL

DE STURENDE FUNCTIE VAN DE DOORSTROOMTOETS

Aan de hand van de resultaten op de doorstroomtoets wordt vastgesteld of leerlingen in voldoende mate de vastgestelde referentieniveaus behalen. In dit artikel betoog ik dat deze gehele systematiek — van het onderwijsresultatenmodel tot de doorstroomtoetsen en de referentieniveaus — in zijn geheel herzien moet worden, omdat de wisselwerking tussen deze drie elementen de kwaliteit van het onderwijs en de kansengelijkheid onder druk zet. Met deze kennis kun je als schoolleider beter duiden waar het landelijke debat over gaat en kun je hier stelling in nemen of meepraten. Daarnaast is het belangrijk te weten wat je als schoolleider wel en niet kan met de data uit de doorstroomtoets, om zo betere beslissingen te maken in het belang van alle leerlingen.

ONDERWIJSRESULTATENMODEL

Het onderwijsresultatenmodel van de Inspectie heeft als doel vast te stellen of leerlingen ‘voldoende’ geleerd hebben, wat wordt gedefinieerd door het behalen van referentieniveaus op de doorstroomtoetsen. De beoordeling hiervan is echter niet gebaseerd op een absolute maatstaf, maar op relatieve signaleringswaarden. De signaleringswaarden zijn zo gekozen dat ongeveer 90% van de scholen resultaten heeft die liggen boven de signaleringswaarden. Er is geen duidelijk vooropgesteld doel; zolang 90% van de

scholen boven een waarde scoort, is er weinig aan de hand.

Het model houdt daarbij ook nog eens de kansongelijkheid in stand door het hanteren van een sterk gedifferentieerde lat op basis van de schoolweging. Bij 1F ligt deze voor alle scholen op een scha-

leerlingen, wat haaks staat op het streven naar gelijke kansen.

Een ander probleem met het onderwijsresultatenmodel is het gegoochel met cijfers. Om tot een eindoordeel te komen, worden de behaalde referentieniveaus van taalverzorging, rekenen en begrijpend lezen bij elkaar opgeteld en gedeeld door

DIT INSTITUTIONALISEERT LAGE VERWACHTINGEN VOOR DE MEEST KWETSBARE LEERLINGEN

mele 85%. Voor 2F/1S is dat anders. Bij scholen met de laagste schoolweging wordt de signaleringswaarde behaald als minder dan 68% van de leerlingen referentieniveau 2F/1S behaalt. Voor scholen met de hoogste schoolweging is dit pas bij minder dan 30% van de leerlingen (Inspectie van het Onderwijs, 2024). In het model wordt er van scholen met een complexe populatie niet verwacht dat leerlingen daar op 2F/1S niveau kunnen lezen en rekenen. De Inspectie moedigt scholen wel aan om hoge ambities te stellen, maar in de praktijk houden scholen hierbij rekening met hun schoolweging en bijbehorende landelijke gemiddeldes. Dat wordt ook zichtbaar in hulpmiddelen die scholen gebruiken om schooleigen doelen te stellen. Dit institutionaliseert lage verwachtingen voor de meest kwetsbare

drie. Dit is een optelsom van ongelijke, niet-vergelijkbare domeinen. Bovendien worden de resultaten van verschillende doorstroomtoetsen – waarvan de vergelijkbaarheid zeer twijfelachtig is – op één hoop gegooid, wat de rekensom nog minder betrouwbaar maakt.

In het onderwijsresultatenmodel worden de resultaten van een school beoordeeld, maar dit is wel een heel smalle blik op onderwijskwaliteit. De onderwijsresultaten van een school worden uitsluitend bepaald aan de hand van één toets in groep 8, die enkel taalverzorging, begrijpend lezen en rekenen meet. De focus op deze beperkte set van meetbare vaardigheden leidt tot ‘teaching to the test’. Door de hoge druk op deze resultaten zien scholen zich genoodzaakt minder waarde te hechten aan vakgebieden die



misschien wel essentieel zijn voor de ontwikkeling van de leerling (zoals geschiedenis, kunst en cultuur, en bewegingsonderwijs), simpelweg omdat deze niet in het model worden meegenomen. De druk die het onderwijsresultatenmodel op scholen legt, werkt hierbij mogelijke fraude in de hand. Voor scholen die onder druk staan (ongeveer 10% van de scholen), kan het verleidelijk zijn om de toets niet strikt volgens protocol af te nemen. Dit kan variëren van het extra aanbieden van specifieke stof vlak voor de toets, het geven van extra tips tijdens de afname, tot, in extreme gevallen, het teruggeven van de nagekeken toetsen aan de kinderen met de opdracht om de fouten nog even te verbeteren. De problemen met het model worden nog

groter als je beseft dat dit gebaseerd is op de resultaten op de doorstroomtoetsen.

DOORSTROOMTOETSEN

In de doorstroomtoets wordt zoals eerder gesteld slechts een beperkt deel van het curriculum getoetst; alleen taal- en rekenvaardigheden worden getoetst. Binnen het domein taal worden alleen lezen en taalverzorging getoetst. Belangrijke vaardigheden als mondelinge taalvaardigheid en schrijven blijven buiten beeld en worden hierdoor in de lespraktijk soms ook als minder belangrijk gezien. Ook voor rekenen kun je je afvragen of alle belangrijke vaardigheden worden getoetst. In de laatste LIB doorstroomtoets staan 60 rekenopgaven. In het referentiekader rekenen staan meer dan drie

keer zoveel doelen (1F en 1S bij elkaar). Het beheersingsniveau van 1F of 1S dat door de doorstroomtoets wordt aangetoond, is indicatief en zeer grofmazig.

Toch wordt het resultaat van de doorstroomtoets tot op de procentpunt nauwkeurig aangegeven. Zoals Van der Steeg (2021) betoogt, is dit merkwaardig, aangezien de grenzen tussen referentieniveaus 1F en 2F/1S soms subjectief zijn en een objectief onderscheid nauwelijks te maken is. De 'precieze' score miskent bovendien de aanwezige betrouwbaarheidsintervallen.

VRAAGSTELLING

Ook op de vraagstelling van de toets is veel af te dingen. Critici stellen dat de vraagstelling in de toetsen de werkelijke vaardigheid niet meet. Eva Naaijken spreekt over 'hoepeltjes' waar leerlingen doorheen moeten leren springen (Ros & Marreveld, 2022). Zo hoeven leerlingen de teksten bij het onderdeel lezen vaak niet werkelijk te doorgronden, maar moeten ze een trucje leren. Effectief, kwalitatief goed leesonderwijs (dat zich richt op diep begrip) wordt daarom door sommige scholen opzij gezet om beter door de 'hoepeltjes' van de toets te springen. En als leerlingen bij taalverzorging gevraagd wordt om fout of goed geschre-

TIPS VOOR DE SCHOOLLEIDER

Als schoolleider heb je zeer beperkte invloed op de verandering van de gehele systematiek op landelijk niveau. Wat kun je op school in ieder geval doen?

1. Stel hoge ambities aan het onderwijs, onafhankelijk van de schoolweging.
2. Stimuleer je team om verder te kijken dan de doorstroomtoets en leerlingvolgsystemen. Betrek altijd meerdere data (bijvoorbeeld methodegebonden toetsen, observaties) voordat er overhaaste conclusies getrokken worden.
3. Bespreek met je team de breedte van het curriculum, dat verder gaat dan alleen de getoetste onderdelen. Bekijk voor taal bijvoorbeeld ook mondelinge taalvaardigheid en schrijven. De nieuwe kerndoelen bieden hiervoor een uitgelezen kans.



ven woorden te herkennen, zie je de discrepantie met effectief spellingsonderwijs (Bosman & Schraven, 2013).

De problemen met de domeinen en vragen van de doorstroomtoetsen worden nog groter gemaakt omdat hetzelfde in beeld moet worden gebracht met verschillende toetsen. Zoals een artikel in de Volkskrant (Meindertsma et al, 2025) pijnlijk duidelijk maakte, maakt het wel degelijk uit welke toets een school kiest. Een leerling die de ene toets maakt, heeft een andere kans om 2F/1S te behalen dan een vergelijkbare leerling die een andere toets maakt. Een toets op papier levert hierbij over het algemeen hogere scores op dan een toets die

BRONNEN

- Bakker et al, (2025). *Reken het rekenonderwijs niet af op percentages leerlingen die 1S behalen*. Volgens Bartjens.
- Bosman, A., & Schraven, J. (2013). *Cito-spellingtoets schaadt het spellingonderwijs*.
- Inspectie van het Onderwijs (2024). *Beoordeling leerresultaten primair onderwijs schooljaar 2024-2025*.
- Meindertsma, B., Frijters, S., & Van den Berg, J. (2025, 27 maart). Het maakt wél uit welke doorstroomtoets een school kiest. *De Volkskrant*.
- Ros, B., & Marreveld, M. (2022). *Eindtoets ongeschikt om leesvaardigheid te toetsen*.
- Van den Broek (2022). *Analyse en evaluatie referentieniveaus Nederlandse taal en rekenen*. Onderzoek in opdracht van het Ministerie van Onderwijs Cultuur en Wetenschap.
- Van Der Steeg, M. (2021). 'Referentieniveaus voor het basisonderwijs'. *Nieuwleren*.

digitaal wordt afgenomen.

REFERENTIELEVELS

De basis onder het problematische onderwijsresultatenmodel en de doorstroomtoetsen wordt gevormd door de referentieniveaus. Het onderwijsveld heeft hier over het algemeen aardig wat vertrouwen in. Dit is echter niet terecht. Uit de eindrapportage referentieniveaus (Van den Broek et al, 2022) blijken er behoorlijk wat problemen te zijn met deze

CRITICI STELLEN DAT DE VRAAGSTELLING IN DE TOETSEN DE WERKELIJKE VAARDIGHEID NIET MEET

referentieniveaus. Zo blijkt dat het Referentiekader Taal op veel fronten tekort schiet. Er is te weinig samenhang tussen domeinen (lezen, schrijven, etc.) en er zijn inconsistenties in benamingen, opbouw van complexiteit en ordening tussen en binnen domeinen. Bovendien zorgt de beschrijving van de taaldomeinen als aparte onderdelen ervoor dat deze in het onderwijs en de toetsing ook als losse domeinen behandeld worden. Dit staat haaks op de behoefte om taalonderwijs juist geïntegreerd aan te bieden; iets wat ook wordt gestimuleerd door de nieuwe kerndoelen.

GEEN IJKING

Bij rekenen-wiskunde heeft er geen ijkingsplaats gevonden om vast te stellen of de inhoudelijke beschrijvingen van 1S daadwerkelijk aansluiten bij wat in het voortgezet onderwijs nodig is (Bakker et al., 2025). Verder is er bij rekenen een fout gemaakt in de wettekst (Van den Broek et al., 2022). De oorspronkelijk doordachte opbouw van de doorlopende leerlijn in twee sporen (F en S) is in de wettekst niet meer goed zichtbaar. Ook zijn er fouten geslopen in het uiteindelijke referentiekader rekenen; alleen de specificaties met voorbeelden zijn overgenomen, zonder de oorspronkelijke inhoudelijke beschrijvingen. Dit kan leiden tot de misvatting dat de voorbeelden doelen op zichzelf zijn. Als gevolg van onduidelijkheden in het Referentiekader hebben scholen soms een vertekend beeld van de na te streven doelen aan het eind van het PO (bij-

voorbeeld 1F voor taal in plaats van 2F, en onduidelijkheid over 1S voor rekenen). Daarnaast verschilt bijvoorbeeld het referentiekader rekenen op onderdelen van andere wettelijke curriculumdocumenten, met name de kerndoelen PO, wat het behoorlijk verwarrend maakt. De referentieniveaus zijn in 2010 vastgesteld en inmiddels niet meer actueel. Actuele wetenschappelijke ontwikkelingen zijn niet of slechts in beperkte mate in het Referentiekader te herken-

nen. Voor taal gaat dit bijvoorbeeld om nieuwe inzichten op het gebied van meertaligheid en digitalisering. Bakker et al. (2025) concluderen dat er diverse goede redenen zijn om de referentieniveaus voor rekenen zelfs af te schaffen.

CONCLUSIE EN VERVOLG

De beoordelingssystematiek voor onderwijsresultaten is een complexe en ondeugdelijke constructie. Het onderwijsresultatenmodel gaat uit van lage verwachtingen, hanteert betekenisloze optelsommen en werkt een ongezonde focus op meetbaarheid en zelfs fraude in de hand. De doorstroomtoetsen, als meetinstrument, zijn onderling onvergelijkbaar en dwingen scholen tot 'teaching to the test'. Tot slot, de referentieniveaus die de basis vormen, zijn inhoudelijk inconsistent, methodologisch gebrekkig en verouderd.

Dit alles overziend zou de eerste stap moeten zijn dat de doorstroomtoets niet meer gebruikt wordt als beoordelingsinstrument voor de onderwijskwaliteit. De doorstroomtoets zou dan alleen een aanvulling moeten zijn op het beeld dat de school heeft van de leerlingen. Daarnaast moet er snel nagedacht worden of en hoe de doorstroomtoets betrouwbaar inzicht kan geven in taal en rekenen en een breder palet aan kennis en vaardigheden. Tenslotte is er met de komst van de nieuwe kerndoelen een kans om het referentiekader opnieuw te bekijken en wellicht te herzien of zelfs af te schaffen.

BSM