



NORMERING VAN EINDTOETSEN VIA EEN GEZAMENLIJK ANKER

Expertgroep Toetsen PO
Ref.nr.: 20.04

Als de normen voor de verschillende eindtoetsen niet vergelijkbaar zijn, worden gelijke prestaties van leerlingen die verschillende eindtoetsen maken niet meer gelijk beoordeeld en wordt het overheidsbeleid om leerlingen gelijke kansen in het onderwijs te geven ondergraven. Om de vergelijkbaarheid van de eindtoetsresultaten beheersbaar te houden, is in 2018 een gezamenlijk anker van opgaven geïntroduceerd. In 2019 is dit anker opnieuw in de eindtoetsen opgenomen en gebruikt voor het vaststellen van de cesuren voor de referentieniveaus en toetsadviezen. In 2020 zijn door de Corona-crisis geen eindtoetsen afgenomen. In 2021 zal dit wel het geval zijn, en ook in dat jaar zal het gezamenlijke anker weer in de eindtoetsen worden opgenomen.

De aanleiding voor het invoeren van een gezamenlijk anker is dat bij de afnamen van de eindtoets in 2015, 2016 en 2017 bij bestudering van de afnamedata aanwijzingen zijn gevonden voor onderlinge verschillen in de toetsadviezen en behaalde referentieniveaus, die niet zijn toe te schrijven aan verschillen in de vaardigheidsniveaus van de populaties van de verschillende eindtoetsen (de DIA-Eindtoets van Diataal B.V., de AMN Eindtoets van AMN, de IEP van Bureau ICE, de Route 8 van A-Vision en de Centrale Eindtoets (CET) ontwikkeld door Stichting Cito en aangeboden onder auspiciën van het CvTE). Deze verschillen zijn gerapporteerd in de Rapportage Vergelijkbaarheid Eindtoetsen (Expertgroep Toetsing PO, December 2016). Hierop heeft het ministerie van OCW de Expertgroep opdracht gegeven om een methode te ontwikkelen en in te voeren om deze ongewenste verschillen te voorkomen.

Om aan deze opdracht gevolg te geven, heeft de Expertgroep Toetsen PO het initiatief genomen om in samenwerking met alle betrokkenen (de verschillende aanbieders van eindtoetsen, het CvTE en het ministerie van OCW) een gezamenlijk anker van opgaven te ontwikkelen. Dit om bij de normering van de eindtoetsen rekening te kunnen houden met de verschillen in de moeilijkheidsgraad van de eindtoetsen en de verschillen in het vaardigheidsniveau van de verschillende populaties van de eindtoetsen. Het gezamenlijk anker wordt ieder jaar deels ververs, zodat de normen, zowel tussen de verschillende eindtoetsen als over de jaren heen, constant kunnen worden gehouden.

In 2018 is het gezamenlijk anker voor de eerste maal in de eindtoetsen opgenomen en toegepast bij de normering. Bij de eerste normering was de situatie nog niet optimaal, omdat de opgaven van het anker nog niet in een high-stakes situatie zoals een eindtoetsafname waren getest. Daardoor waren hun psychometrische eigenschappen, zoals de moeilijkheidsgraad, nog niet exact bekend. Ondanks dat waren de resultaten vertrouwenwekkend en secundaire analyses na de afname en normering toonden aan dat de ankeropgaven voor het vergelijken van de eindtoetsen goed hadden gefunctioneerd. Ook bleken er geen significante verschillen tussen traditionele en digitale afnamen en eventueel adaptieve afnamen (voor details zie de rapportage Vergelijkbaarheid

eindtoetsen en invoering van gezamenlijk anker, 2018, Expertgroep Toetsing PO, Ref.nr. 18.01). In de afname van 2019 en toekomstige afnamen is het probleem van het schatten van de moeilijkheidsgraad van ankeropgaven overigens opgelost, omdat alle toetsaanbieders de potentiële ankeropgaven wel in een high-stakes situatie testen, door ze als niet-meetellende opgaven in hun eindtoets op te nemen.

In 2019 is de procedure van normering via het gezamenlijk anker opnieuw uitgevoerd. Initieel ontstond bij het vaststellen van cesuren voor de toetsadviezen het probleem dat het aantal adviescategorieën was teruggebracht van 10 categorieën in 2018 naar 6 categorieën in 2019. Nadat dit probleem was opgelost, kon een normering worden vastgesteld die voor alle betrokken (toetsaanbieders, scholen, CvTE, Expertgroep en OCW) acceptabel was. (voor details zie de rapportage Normering afname eindtoets 2019, Expertgroep Toetsing POP, ref. nr. 19.11).

Uit secundaire analyses na de eindtoetsafname bleek opnieuw dat de ankeropgaven goed hadden gefunctioneerd en dat verschillen tussen de moeilijkheidsgraad van de verschillende eindtoetsen en de populaties van de verschillende eindtoetsen nauwkeurig in kaart waren gebracht. Daardoor waren alle toetsadviezen goed vergelijkbaar. Deze vergelijkbaarheid geldt uiteraard gegeven de vrijheid die toetsaanbieders hebben om hun eindtoets in te vullen, zoals de mogelijkheid van verschillende extra onderdelen naast de referentieonderdelen op te nemen. Deze vrijheid om de eindtoets in te vullen hebben de eindtoetsaanbieders; in de procedure wordt er wel voor gewaakt dat dit niet leidt tot systematische bevoordeling van groepen leerlingen.

Bij de afname in 2021 wordt de procedure voor normering via een gezamenlijk anker opnieuw uitgevoerd. Met invoering van de nieuwe wetgeving in 2022 komt de regie over de kwaliteitsbewaking van de eindtoetsen en hun vergelijkbaarheid bij het CvTE te liggen. Op dit moment wordt door de Expertgroep Toetsing PO en het CvTE nauw samengewerkt ten behoeve van een zorgvuldige en naadloze overdracht van ontwikkelde procedures. In nauw overleg met alle betrokkenen zijn de procedures verder verfijnd. Zo zal bij de vaststelling van de normen structureel ook informatie over de schooladviezen beschikbaar zijn, zodat de kwaliteit van de normering via het gezamenlijk anker nog beter kan worden gegarandeerd en daarmee de vergelijkbaarheid van de verschillende eindtoetsen wordt versterkt.