

Ministerie van Onderwijs, Cultuur en Wetenschap
T.a.v. de heer S. Dekker, staatsecretaris van Onderwijs, Cultuur en Wetenschap
Postbus 16375
2500 BJ Den Haag

Amsterdam, 10 december 2013

Betreft: Internetconsultatie Toetsbesluit PO

Geachte heer Dekker,

Graag maken wij, uitgevers van test- en toetsmateriaal voor het primair en voortgezet onderwijs, gebruik van de internetconsultatie betreffende het Toetsbesluit PO.

We rubriceren onze opmerkingen per artikel van het uitvoeringsbesluit. In de nota van toelichting wordt aangegeven dat sommige onderdelen van het Toetsbesluit PO nog nader dienen te worden ingevuld. Het betreft onder andere het beoordelingskader van de expertgroep die de minister zal adviseren over toelating van andere eindtoetsen en toetsen in het kader van het leerlingvolgsysteem. Een aantal van onze opmerkingen kunnen als prematuur worden beschouwd omdat ze met name betrekking hebben op deze uitwerking van het uitvoeringsbesluit, waarvan alleen de contouren in het Toetsbesluit PO zijn beschreven. Het leek ons desalniettemin nuttig ze hier al te melden.

Betreffende artikel 4. Kenmerken eindtoets

Ad b)

Naast de eisen van validiteit en betrouwbaarheid wordt 'een deugdelijke normering' van de eindtoets vereist. In principe geschiedt de normering van de toets onder dezelfde gestandaardiseerde voorwaarden als bij de eigenlijke afname van een toets. Een eindtoets wijkt hiervan in zoverre vanaf dat deze eenmalig wordt afgenomen (bij het adaptief toetsen is sprake van een uitzonderlijke situatie) waarna de normering wordt vastgesteld. Normering vooraf stuit op praktische bezwaren, gezien de wenselijkheid van geheimhouding van de inhoud van de toets. De gebruikelijke procedure bij de constructie van een eindtoets is kalibratie van de opgaven, waarna een initiële (representatieve) normering kan worden bepaald, welke pas na de afname definitief wordt vastgesteld. Dit betekent dat bijvoorbeeld de reken- en de taalopgaven kunnen worden gekalibreerd en genormeerd op gescheiden representatieve steekproeven, en de opgaven in deze fase bijvoorbeeld niet noodzakelijk in dezelfde volgorde worden aangeboden als bij de eigenlijke afname.

Met andere woorden, we verwachten dat de expertgroep in haar beoordeling van de deugdelijkheid van de normering rekening houdt met de bijzondere aspecten in de uitvoering die een normering van een eindtoets vraagt.

Ad c)

De inhoud van de eindtoets is gebaseerd op de toetswijzers van de CvE; maar dat laat onverlet dat een eindtoets onderdelen kan bevatten, die niet expliciet in de toetswijzers worden uitgewerkt. Een voorbeeld zou Begrijpend Luisteren kunnen zijn, een onderdeel dat in de huidige eindtoets niet wordt getoetst, maar wel in de referentieniveaus is opgenomen.

Ad d)

In dit artikel wordt aangegeven dat de opgaven van de eindtoets jaarlijks worden *ververst*. In de nota van toelichting staat dat de eindtoets 'jaarlijks bestaat uit *nieuwe opgaven* voor taal en rekenen' (p. 10). Op p. 11 van de toelichting wordt gesteld dat terugkerende, reeds gebruikte opgaven een bedreiging vormen voor de betrouwbaarheid. Dit is op zichzelf correct, maar het gevaar wordt veel kleiner wanneer het om slechts een deel van de opgaven gaat. In uw brief van 26 november 2013 aan de

Tweede Kamer over de uitwerking van andere eindtoetsen stelt u onder de inhoudelijke voorwaarden: 'de toets wordt jaarlijks voldoende ververs't. Het zou wenselijk zijn deze nuancering ook in de AmvB op te nemen, waarbij zij opgemerkt dat aanvulling van de toets met extra opgaven in het geval van digitale toetsing ook als verversing cq. vernieuwing moet worden gezien.

Ad e)

De eindtoets dient inzicht te geven in de mate van beheersing van de referentieniveaus. In de nota van toelichting wordt ook het domein taalverzorging genoemd (p. 12). De referentieset om een eindtoets ook te kunnen verankeren op het domein taalverzorging is echter voorlopig nog niet beschikbaar en daardoor ook niet uitvoerbaar. Indien mogelijk graag in de toelichting aangeven wanneer dat naar verwachting wel mogelijk is.

Ook wanneer er voor een te toetsen onderwerp geen referentieset beschikbaar is, moet het mogelijk zijn dit in de eindtoets op te nemen.

Een ander punt van zorg is de te onderscheiden domeinen bij Rekenen-Wiskunde; per domein is de vaardigheidsschaal qua moeilijkheidsgraad ongelijk verdeeld. Bij adaptief toetsen zijn de items die verschillende domeinen representeren ook om praktische redenen niet gelijk verdeeld; de items op het onderdeel Verbanden (tabellen, grafieken, et cetera) zijn vaak relatief ondervertegenwoordigd omdat dit onderdeel relatief tijdsintensief is om te toetsen en de opgaven in het hogere gebied van de vaardigheidsschaal zijn geplaatst, vergeleken met bijvoorbeeld het onderdeel Getallen (vermenigvuldigen, breuken, et cetera). Wanneer echter eenzelfde meetprecisie zou worden vereist voor alle domeinen, zou dit betekenen dat de toets inhoudelijk onevenwichtig wordt, omdat de leerlingen relatief veel tijd kwijt zouden zijn met het maken van de opgaven betreffende het domein Verbanden, en voor sommige leerlingen met een laag vaardigheidsniveau, bij voorbeeld lwooleerlingen, is het bovendien de vraag of er sowieso voor hen opgaven kunnen worden geconstrueerd in dit domein die deze leerlingen kunnen maken.

Om de eis van aparte rapportage met subschalen voor de vier rekendomeinen te relativieren, is ook een sterk psychometrisch argument te geven; in de huidige Cito Eindtoets laten de verschillende domeinen een correlatie zien van boven de .97 (Boxtel, et. al., 2010, p. 97). Voor zover deze zich al voordoen, zullen eventuele niveauverschillen op de vier rekendomeinen waarschijnlijk voornamelijk een meetfout representeren.

Ad g)

De eindtoets heeft onder andere de functie voor scholen om zich te kunnen verantwoorden naar de Inspectie voor hun behaalde eindopbrengsten. In de huidige praktijk vereist de Inspectie hiervoor een representatieve steekproef, eventueel over opeenvolgende schooljaren gestapeld, van circa 150 scholen. Daaraan kan bij een alternatieve eindtoets moeilijk worden voldaan. Om het toch mogelijk te maken dat een eindtoets kan worden gebruikt voor verantwoording van de eindopbrengsten kan worden volstaan met een kleinere steekproef of als criterium worden gesteld het percentage leerlingen dat het fundamenteel niveau haalt op de onderdelen rekenen en taal. De gestelde eisen aan een eindtoets maakt dit immers mogelijk. Het is noodzakelijk voor de realisatie van het beoogde gelijke speelveld dat in de toelichting in dit opzicht een voorbehoud wordt gemaakt.

In de Toelichting wordt aangegeven dat de huidige mogelijkheden 'om de meting te corrigeren' blijven bestaan. We veronderstellen dat hiermee de verschillende normeringen worden bedoeld, die de Inspectie hanteert, voor scholen met relatief weinig achterstandsleerlingen en scholen met relatief veel achterstandsleerlingen.

Betreffende artikel 9. Procedure toelating andere eindtoetsen

Ad 1)

Voor betreffende de normering van de eindtoets, zie onze eerdere opmerking bij artikel 4 ad b). Wat de rapportage van de beheersing van de referentieniveau betreft, dit is uiteraard alleen mogelijk indien het CvE zowel de referentieset als de te hanteren cesuur binnen redelijke termijn ter beschikking heeft gesteld.

Betreffende de mogelijkheid van een tweede afnamemoment van de toets, geldt dezelfde voorwaarde als vermeld onder artikel 4 ad d); het gaat dan niet om een exacte kopie van de eerder dat jaar afgenomen eindtoets, maar een toets waarbij een deel van de opgaven zijn vervangen.

Betreffende artikel 11. Kwaliteit toetsen

Ad 2)

Het kwaliteitsoordeel dat door de expertgroep wordt gegeven over een enkele toets of over een reeks van toetsen, zou o.i. uitdrukkelijk alleen betrekking moeten hebben op nieuwe toetsen; toetsen die reeds door de COTAN als 'voldoende' zijn beoordeeld, zouden niet opnieuw door de expertgroep moeten worden beoordeeld en moeten door scholen gebruikt kunnen blijven worden voor hun leerlingvolgsystemen. Dit i.v.m. de noodzaak van een gelijk speelveld.

Ad 3a)

Aan een leerlingvolgsysteem wordt de eis gesteld dat de leervorderingen van de leerlingen systematisch worden gemeten. In de huidige praktijk bestaan hiervoor twee methoden, namelijk vaardigheidsscores en DLE's. De DLE's die in het leerlingvolgsysteem worden gehanteerd zijn vaak omgezette vaardigheidsscores, maar niet altijd. Het Cito heeft bezwaren tegen het gebruik van DLE's, waarachter de COTAN zich heeft geschaard, maar deze zijn ook weersproken. Zonder hier op deze discussie in te willen gaan, is het een feit dat DLE's breed zijn verankerd in de leerlingvolgsystemen. Over de vergelijkbaarheid van de resultaten wordt in de toelichting opgemerkt dat de leerlingsscores van een leerlingvolgsysteem:

"te plaatsen zijn op een zogenoemde vaardigheidsschaal. Een vaardigheidsschaal richt zich op bepaalde kennis of vaardigheden, zoals de woordenschat van leerlingen (-) De losse toetsen van het leerling- en onderwijsvolgsysteem meten steeds een volgend deel van de vaardigheidsschaal om zo de vooruitgang oftewel leerwinst van de leerling op deze kennis of vaardigheid gedurende een bepaalde tijd vast te stellen. "

De term 'vaardigheidsschaal' is een in het Cito LOVS gebruikte term. Om misverstanden te vermijden zou het wenselijk zijn wanneer in de toelichting hieraan zou worden toegevoegd 'uit te drukken in vaardigheidsscores of DLE's'.

Ad 3b)

Toetsen die in het leerlingvolgsysteem worden ingezet, dienen ook te worden gebruikt voor de vergelijkbaarheid van de resultaten op schoolniveau. De vraag is in hoeverre hiervoor een extra normering op schoolniveau wordt vereist, naast een representatieve normering op leerlingniveau. Zie hiervoor opnieuw onze opmerking bij artikel 4 ad d).

Betreffende het amendement Straus. Kwaliteit van leerling- en onderwijsvolgsystemen

Het amendement Straus heeft een aantal extra eisen gesteld aan een leerling- en onderwijsvolgsysteem, naast de eerder genoemde eisen van inhoudelijke validiteit, betrouwbaarheid en deugdelijk normering. Het is de bedoeling dat de expertgroep ook deze eisen in haar oordeel betreft.

Het probleem is echter dat laatstgenoemde eisen strikt gesproken van toepassing zijn op toetsen, zoals ook de COTAN deze in haar beoordelingen gebruikt, terwijl de eisen die in het amendement Straus worden geformuleerd, deze lijken te overschrijden. Met andere woorden, hier lijkt de functie van een leerling*administratie*systeem en een leerling*volg*systeem te worden verward. Deze vallen soms samen, maar niet altijd. Zo kan een uitgever een leerlingvolgsysteem ontwikkelen, dat wil zeggen een

toetsbatterij om de cognitieve voortgang van leerlingen te volgen, zonder zich te bemoeien met de administratieve registratie van de toetsresultaten (Esis, Parnassys, dot.com) en visa versa. In de huidige praktijk hebben de volgende eisen eerder betrekking op het leerlingadministratiesysteem dat de scholen hanteren, en niet het leerlingvolgsysteem (Toelichting, p. 19):

- ieder leerling- en onderwijsvolgsysteem beschikt over een functionaliteit waarmee de leervorderingen inzichtelijk kunnen worden gemaakt voor ouders;
- een leerling- en onderwijsvolgsysteem moet zo min mogelijk extra bureaucratie opleveren;
- een leerling- en onderwijsvolgsysteem kent een exporteerfunctie voor data-uitwisseling met het voortgezet onderwijs;
- de gegevens uit het leerling- en onderwijsvolgsysteem zijn in eigendom van de school en niet in eigendom van de aanbieder van het leerling- en onderwijsvolgsysteem;
- ieder leerling- en onderwijsvolgsysteem waarborgt de privacy van de leerling.

Mogelijk dat de expertgroep op deze aspecten de huidige leerlingadministratiesystemen kan beoordelen, zonder deze eisen direct aan een toets of een reeks van toetsen te verbinden.

Dit zijn onze voornaamste bedenkingen bij het Toetsbesluit PO. Uiteraard zijn wij graag bereid onze deze bedenkingen toe te lichten.

Hoogachtend,

Hilde Hacquebord, Diataal/Rijksuniversiteit Groningen
Ton Kapinga, 678 Onderwijs Advisering
Hermien Lubbers, A-Vision Advisering en Begeleiding B.V.
John Mulder, AMN Systems BV
Judith Rood, Boom test uitgevers
Gert Jaap Schoppink, Pearson Assesement and Information BV
Franke Teunisse, Bureau ICE
Klaas ter Veen, Projectgroep Schooleindonderzoek