



Regeling van het College voor toetsen en examens van 1 december 2025, nummer CvTE-25.012.15, houdende wijziging van de Regeling beoordelingsnormen doorstroomtoetsen PO in verband met actualisering van de bijlage en kleine verbeteringen

Het College voor toetsen en examens,

Gelet op artikel 3a, eerste lid, onderdeel f, van de Wet College voor toetsen en examens;

Gezien de goedkeuring van de Staatssecretaris van Onderwijs, Cultuur en Wetenschap, gegeven op 15 december 2025, nummer 1783366,

Besluit:

ARTIKEL I

De bijlage bij de Regeling beoordelingsnormen doorstroomtoetsen PO wordt vervangen door de bijlage bij deze regeling.

ARTIKEL II. INWERKINGTREDING

Deze regeling treedt in werking met ingang van de dag na de datum van uitgifte van de Staatscourant waarin de regeling wordt geplaatst.

Deze regeling zal met de toelichting in de Staatscourant worden geplaatst.

*Het College voor toetsen en examens,
de voorzitter,
J.H. van der Vegt*



BIJLAGE ALS BEDOELD IN ARTIKEL I

BIJLAGE 1. DE PROCEDURE OM TE KOMEN TOT DE BEOORDELINGSNORMEN VAN DE DOORSTROOMTOETSEN

Bijlage behorende bij artikel 1 van de Regeling beoordelingsnormen doorstroomtoetsen PO

1.1. Reikwijdte

De Regeling beoordelingsnormen doorstroomtoetsen po bevat algemeen verbindende voorschriften voor toetsaanbieders van een tot het stelsel toegelaten doorstroomtoets po. De Regeling beoordelingskader doorstroomtoetsen po schrijft voor dat een toets alleen zal worden erkend als de procedure om te komen tot de beoordelingsnormen van de doorstroomtoets wordt gevolgd.

De Regeling Beoordelingsnormen Doorstroomtoetsen po beschrijft de inhoud en de processtappen van de volgende sleutelonderdelen, met verwijzing naar de paragrafen met kwaliteitseisen zoals opgenomen in de Regeling beoordelingskader doorstroomtoetsen po:

Ter voorbereiding op de landelijke (toetsoverstijgende) normering:

- Gezamenlijk anker

De landelijke (toetsoverstijgende) normering (conform deze onderhavige Regeling beoordelingsnormen doorstroomtoetsen po).

Na afronding van de landelijke (toetsoverstijgende) normering:

- Datalevering toelatings- en doorstroomonderzoek

Een doorstroomtoets wordt erkend door het CvTE voor een periode van vier jaar, mits deze voldoet aan de voorwaarden voor erkenning en de regeling beoordelingskader doorstroomtoetsen po.

Jaarlijks zal het CvTE vaststellen of de in dat schooljaar aan te bieden erkende doorstroomtoets voldoet aan de criteria op basis waarvan de erkenning is verleend. De erkende doorstroomtoets is onderdeel van de jaarlijks door een adviseur uit te voeren landelijke normeringsprocedure zoals beschreven in deze Regeling Beoordelingsnormen. Als een doorstroomtoets niet (langer) aan de voorschriften van deze Regeling Beoordelingsnormen voldoet, kan de erkenning door het CvTE worden ingetrokken op grond van artikel 3a, vierde lid, van de Wet College voor toetsen en examens.

1.2. Begrippen en definities

<i>Anker</i>	Het anker bestaat uit opgaven die in twee of meer toetsen worden opgenomen om verschillen in de moeilijkheidsgraad tussen de betreffende toetsen vast te kunnen stellen
<i>Ankerbank</i>	De bank met vastgestelde gezamenlijk ankeropgaven.
<i>Ankeropgave</i>	Een opgave in het anker
<i>Aspirant-ankeropgave</i>	Een opgave die mogelijk in het gezamenlijk anker opgenomen kan worden, maar nog niet als zodanig is vastgesteld.
<i>Cesuur</i>	Minimale prestatie (gerepresenteerd als vaardigheid of toetsscore) die net indicatief is voor het halen van een (referentie)niveau of toetsadviescategorie.
<i>Equivaleren</i>	verschillende doorstroomtoetsen op dezelfde schaal brengen.
<i>Jaarset</i>	De selectie van gezamenlijk ankeropgaven uit de ankerbank die in een bepaald afnamejaar wordt afgenomen.
<i>Kindopgave</i>	Aangepaste versie van een ankeropgave van een toetsaanbieder om de opgave te laten aansluiten bij een specifieke doorstroomtoets.
<i>Moederopgave</i>	Basisversie van een ankeropgave in de ankerbank. Een aspirant-ankeropgave kan een moederopgave worden als het is gepretest en initiële psychometrische parameters bekend zijn, en de opgave inhoudelijk en psychometrisch voldoet. Eventuele kindopgaven zijn afgeleid van deze versie.
<i>Pretest</i>	Hier wordt dezelfde definitie toegepast als in het beoordelingskader, wat betekent dat dit zowel afnamen los van de operationele afname kan betreffen, alsook de afname in de high-stakes doorstroomtoets met beperkte exposure (zaaien).
<i>De adviseur</i>	Het onderdeel van Stichting Cito waar de wettelijke adviestaak voor het CvTE belegd is, de adviseur.
<i>Toetsaanbieder</i>	Aanbieder van een doorstroomtoets.
<i>Werkgroep gezamenlijk anker</i>	Een werkgroep waar alle toetsaanbieders in het stelsel vertegenwoordigd zijn.

1.3. Eisen aan de software

De functionaliteit van de software dient passend en toereikend te zijn voor het uitvoeren van de landelijke (toetsoverstijgende) normeringsprocedure, voor gebruik van een twee parameter logistisch itemresponsemodel (2PL) – dat wil zeggen een model waarin de moeilijkheidsgraad en het onderscheidend vermogen van de toetsopgaven worden meegenomen.



Het CvTE laat periodiek de software, die door de adviseur wordt gebruikt voor de landelijke toetsoverstijgende normering, extern valideren op bovenstaande criteria.

2. Gezamenlijk anker

Het gezamenlijk anker is een verzameling van opgaven die in alle doorstroomtoetsen wordt opgenomen om verschillen in de moeilijkheidsgraad tussen de betreffende doorstroomtoetsen vast te kunnen stellen en daarmee zowel de vergelijkbaarheid tussen doorstroomtoetsen als over jaren heen te waarborgen. Het ankerbankbeheer is belegd bij de adviseur.

2.1. Uitgangspunten gezamenlijk anker

- Het gezamenlijke anker is bij wet een instrument waar alle toetsaanbieders aan mogen bijdragen door het leveren van opgaven.
- Het gezamenlijk anker wordt gebruikt om de toetsadviezen en de behaalde referentieniveaus voor Lezen, Taalverzorging en Rekenen van de verschillende doorstroomtoetsen te equivaleren.
- Om de vergelijkbaarheid tussen de toetsen te vergroten is het belangrijk dat al vanaf een vroeg stadium in de opgaveconstructie er rekening mee wordt gehouden dat de opgaven met minimale aanpassingen weergegeven kunnen worden in alle toetsplatformen, en dat de opgaven inhoudelijk bij de reguliere opgaven van alle toetsen aansluiten.
- Het gezamenlijk anker bevat opgaven die afneembaar zijn en onderwijskundig en psychometrisch passen bij alle doorstroomtoetsen.
- Het gezamenlijk anker voor een bepaald afnamejaar is gebaseerd op het voor dat jaar geldende beoordelingskader, toetswijzer en toetsbesluit¹.
- De toetsmatrijs is de basis voor het ontwikkelen van nieuwe gezamenlijk ankeropgaven en het samenstellen van de jaarset (zie paragraaf 2.4).
- De toetsaanbieders zorgen voor geheimhouding van de gezamenlijk ankeropgaven. De inhoud van gezamenlijk ankeropgaven wordt alleen gedeeld via een speciaal hiervoor opgezette server, en dus niet via e-mail of online vergaderingen. Gecontroleerde exposure van de opgaven draagt ook bij aan de geheimhouding.
- De toetsaanbieders accepteren ieder afnamejaar de jaarset van gezamenlijk ankeropgaven (zie ook 2.2.1D) en nemen dit **volledig** (d.w.z.: alle opgaven) op in hun toets.
- De jaarsets van het gezamenlijk anker worden jaarlijks vernieuwd volgens de hieronder beschreven verversingsstrategie (zie 2.2.1F).
- Het gezamenlijk anker wordt in iedere doorstroomtoets op een manier afgenomen zodat de prestatie op deze opgaven de potentie van leerlingen reflecteert.
- Er bestaat een link tussen de verschillende jaarsets van het gezamenlijk anker. Dat betekent dat er in de jaarset ook ankeropgaven opgenomen zijn die in een eerdere afname gebruikt zijn.
- Het CvTE stelt de jaarset van het gezamenlijk anker vast. Een jaarset kan daarna niet meer gewijzigd worden, tenzij na afname van het er aan voorafgaande jaar een opgave psychometrisch niet blijkt te functioneren als anker. De adviescommissie verifieert bij de vaststelling of het gezamenlijk anker voldoet aan de eisen zoals geformuleerd in deze regeling (zoals samenstelling e.d.).

2.2. De ankercyclus

Deze paragraaf beschrijft de levenscyclus van een ankeropgave. Per stap wordt beschreven wat de te nemen acties zijn, wie wat oplevert, en hoe de verantwoordelijkheden zijn belegd.

2.2.1. Stappen in de ankercyclus

A. Inventarisatie van de ankerbank

Er wordt een inventarisatie gemaakt van de hoeveelheid beschikbaar materiaal in de huidige ankerbank in relatie tot de toetsmatrijs (zie paragraaf 2.4) en de verversingsstrategie (zie stap F). Op basis hiervan wordt door het CvTE aan de ankerbankbeheerder de opdracht gegeven tot het opleveren van een nieuwe jaarset, waar afhankelijk van de inventarisatie een constructieopdracht aan verbonden is. Wanneer er een constructieopdracht nodig blijkt om de jaarset te kunnen leveren, dan zal deze met de toetsaanbieders gedeeld worden.

¹ Voor de meest recente publicaties zie www.toetsenpo.nl



B. Constructieopdracht

De constructieopdracht geeft aan wat voor opgaven opgeleverd moeten worden aan de ankerbankbeheerder en aan welke eisen deze moeten voldoen.

C. Constructie en selectie

De aspirant-ankeropgaven doorlopen in principe hetzelfde traject als normale opgaven voor doorstroomtoetsen: ze moeten voldoen aan dezelfde eisen als reguliere opgaven zoals beschreven in het beoordelingskader.

Opgaven dienen daarom ook gepretest te worden voordat ze opgenomen kunnen worden in de ankerbank. Hier zijn twee routes voor:

- A. Toetsaanbieders leveren opgavemateriaal aan, wat vervolgens in de werkgroep gezamenlijk anker besproken wordt en indien nodig aangepast. Geselecteerde opgaven worden vervolgens in verhouding tot het aantal ingeschreven leerlingen verdeeld over de toetsen om gepretest te worden. Als de opgaven na de pretest aan de psychometrische eisen blijken te voldoen, kunnen de opgaven opgenomen worden in de ankerbank.
- B. Toetsaanbieders leveren reeds gepreteste opgaven aan. Dit kan uit hun reeds bestaande bank, of gericht geconstrueerd volgens de opdracht. De opgaven worden besproken in de werkgroep gezamenlijk anker en eventueel geselecteerd om in de ankerbank te worden opgenomen. Eventueel kunnen inhoudelijke aanpassingen voorgesteld worden. Bij grote aanpassingen zal het CvTE besluiten dat een opgave eerst opnieuw in een pretest moet worden opgenomen.

Voordat opgaven gepretest worden in optie A of na een voorgestelde wijziging in route B, dienen de opgaven te zijn voorgelegd en als goed te zijn beoordeeld door de onderwijskundige beoordeling van het CvTE.

Naast de inhoudelijke eisen zoals gespecificeerd in de constructieopdracht moeten de opgaven aan psychometrische eisen voldoen (zie paragraaf 2.3.2).

D. Definitieve selectie en vaststelling jaarset gezamenlijk anker

Uit de ankerbank wordt een selectie gemaakt voor een jaarset met daarin de gezamenlijk ankeropgaven die worden afgenomen in een bepaald afnamejaar. Deze wordt vastgesteld door het CvTE. De in de bank opgenomen *moederopgaven* worden vervolgens door de toetsaanbieders overgenomen en eventueel aangepast zodat ze geschikt zijn voor hun doorstroomtoets. Hierbij worden de richtlijnen uit paragraaf 2.3.4 in acht genomen om mogelijke bias te voorkomen. Aangepaste opgaven of *kindopgaven* worden ter goedkeuring schriftelijk voorgelegd aan het CvTE. Goedgekeurde kindopgaven kunnen worden opgenomen in de doorstroomtoetsen.

E. Afname en analyse

Nadat de doorstroomtoetsen afgenomen zijn, wordt tijdens de kalibratie ten behoeve van de landelijke normering gecontroleerd of de ankeropgaven voldoen aan de psychometrische eisen van een goed anker. Als dit het geval is, worden ze ook daadwerkelijk gebruikt voor de equivalering van de verschillende doorstroomtoetsen en afnamejaren. Indien een ankeropgave psychometrisch voldoet, zal deze ook meetellen in de totaalscore van de leerling. Wanneer een ankeropgave in de afname *niet* blijkt te voldoen, kan dit afhankelijk van de details verschillende gevolgen hebben. Als een opgave niet blijkt te functioneren afname in alle toetsen of toetsvarianten kan de ankeropgave compleet verworpen worden, waardoor de opgave niet meer meewerkt in de equivalering en ook niet meer meetelt in de score van de leerlingen. In het geval van toets-DIF kan er bijvoorbeeld ook gekozen worden om de ankeropgave bij één toets(variant) of bij een deel van de toetsen los te koppelen, waardoor deze opgave voor deze toets(variant) geen rol speelt in de equivalering, maar de opgave wel gebruikt kan worden in de vaardigheidsschattingen en scoring van de leerlingen.

Afhankelijk van de analyse kan een opgave vervolgens aangepast worden voor een volgende afname, of wordt de opgave compleet verwijderd uit de ankerbank. Naast toets-DIF kan er tussen afnamejaren ook parameterdrift plaatsvinden bij een ankeropgave. In dit geval kunnen soortgelijke maatregelen getroffen worden.

F. Verversingsstrategie

Er wordt een verversingsstrategie van 50% nagestreefd. Dit betekent dat iedere jaarset idealiter voor ieder vaardigheidsdomein voor de helft uit opgaven bestaat die voor het eerst ingezet worden. Een kwart van de jaarset bestaat uit opgaven die een jaar eerder in de jaarset opgenomen waren, en een

kwart uit opgaven die voor het laatst meer dan een jaar geleden geselecteerd zijn.

	Fractie opgaven in jaarset jaar X
Nieuwe opgaven	50%
Opgaven uit jaarset jaar X-1	25%
Opgaven uit jaarsets van jaar X-2 of eerder	25%

Nadat een opgave in 3 jaarsets is opgenomen, wordt de opgave niet meer geselecteerd. Een opgave mag maximaal 5 jaar na de eerste opname in een jaarset nog geselecteerd worden.

2.3. Richtlijnen

2.3.1. Richtlijnen rond de constructie van ankeropgaven

- Ankeropgaven moeten voldoen aan de kwaliteitscriteria V2 en V3 van het beoordelingskader.
- Ankeropgaven kunnen met minimale aanpassingen bij alle toetsaanbieders afgenomen worden conform de richtlijnen in paragraaf 2.3.4 (richtlijnen afdruk).
- Ankeropgaven zijn lokaal onafhankelijk.
- In alle ankeropgaven wordt gebruik gemaakt van interpunctie (starten met hoofdletter, eindigen met een punt).
- Geen sleepvragen, cloze-items of opgaven met “aanklikken van zinsdelen”.
- Meerkeuze-opgaven voldoen aan de eisen uit de Checklist voor het beoordelen van de kwaliteit van observatiecategorieën en toetsopgaven (CvTE, 2024).
- Richtlijnen voor opgaven bij het onderdeel Lezen:
 - Altijd twee of meer opgaven per tekst
 - Bij voorkeur opgaven van twee of meer taakuitvoeringen per tekst
- Richtlijnen voor opgaven bij het subdomein Taalverzorging:
 - Voor het onderdeel spelling niet-werkwoorden bestaat een set dictee-opgaven en een set niet-dictee-opgaven.

2.3.2. Richtlijnen rond de selectie van ankeropgaven

Deze richtlijnen betreffen de selectie van de moederopgaven waarvoor in een pretest gegevens zijn verzameld. Wanneer deze opgaven geselecteerd worden, kan iedere toetsaanbieder een eigen kindopgave van deze opgave maken dat zo min mogelijk afwijkt van de moederopgave, maar past bij de eigen vormgeving. Ingediende opgaven hebben niet eerder in een doorstroomtoets van een aanbieder gezeten, tenzij dit in een zaaisteekproef was zoals omschreven in 5.2 van de Regeling beoordelingskader doorstroomtoets PO.

- Op basis van pretestgegevens voldoen de ankeropgaven in ieder geval aan de volgende eisen:
 - P-waarde is lager dan 0,80 en groter dan 0,30².
 - Opgaven hebben een rir groter dan 0,20

Als de steekproefpopulatie van de pretest sterk afwijkt van de toetspopulatie kan een schatting van de p-waarde en rir gemaakt worden door middel van een simulatie.

- Binnen een jaarset van een vaardigheidsdomein hebben de opgaven voldoende spreiding in moeilijkheid zodat er betrouwbaar in een breed vaardigheidsgebied gemeten kan worden, en er moet voldoende spreiding in de inhoudelijke kenmerken zitten.
- Wanneer reeds gepreteste aspirant-ankeropgaven aangeboden worden aan de werkgroep gezamenlijk anker moet bekend zijn hoe vaak en met hoeveel observaties deze opgaven eerder door de toetsaanbieder zijn afgenomen in een pretestsituatie.
- Leesteksten worden niet meer dan twee jaar achter elkaar geselecteerd in een jaarset.

2.3.3. Richtlijnen rond de opslag van ankeropgaven

- Labels worden toegekend door de ankerbankbeheerder en dienen eenduidig te zijn en niet veranderd te worden. Deze behoren geen informatie te bevatten (zoals een afnamejaar) en functioneren alleen ter identificering van de opgave.
 - Rekenen: AR0001, AR0002, enzovoort.
 - Taal: AT0001, AT0002, enzovoort.
 - Lezen: ALAA01, ALAA02, ALAB01, ALAB02, enzovoort.

² Vooral bij Lezen, teksten, is het niet altijd haalbaar om hier precies aan te houden. Ook hier geldt dat dit een richtlijn is en pragmatisch en weloverwogen een keuze wordt gemaakt.

Twee letters na AL geven de tekst aan, numerieke sortering moet gelijk zijn aan de volgorde van de opgaven bij een tekst.

- Er is een eenduidige omschrijving van typering/kenmerken van opgaven, gebaseerd op de toetsmatrijs.
- Van iedere opgave worden op een centrale server de volgende gegevens bijgehouden:
 - De moederopgave in tekstvorm en eventuele bijbehorende afbeeldingen in hoge resolutie.
 - Afbeeldingen van kindopgaven van alle toetsaanbieders en afnamejaren.
 - Klassieke psychometrische gegevens zoals bekend uit de pretest.
- Per toetsaanbieder en afnamejaar worden klassieke psychometrische gegevens en IRT-kenmerken bijgehouden.
- De ankerbank wordt regelmatig gecontroleerd op actualiteit door de ankerbeheerder in samenwerking met de werkgroep gezamenlijk anker, eventueel voortschrijdend inhoudelijk inzicht of veranderingen in technische mogelijkheden.

2.3.4. Richtlijnen rond de afname van het anker in de doorstroomtoets

- De jaarset van het gezamenlijk anker kan integraal worden opgenomen in een aparte toetsvariant of in deelsets verdeeld zijn over verschillende afnamevarianten. Het is daarbij een vereiste dat de verschillende afnames van het gezamenlijk anker bij een toetsaanbieder ook via een intern anker zijn gekoppeld aan de rest van de afname(varianten) zodat er een verbonden design ontstaat wat met de gefixeerde jaarset van het gezamenlijk anker gekalibreerd kan worden.
- Het gezamenlijk anker wordt voorgelegd aan een steekproef van leerlingen uit het regulier basisonderwijs die representatief is voor de toetsaanbieder van de doorstroomtoets. In het geval dat deelsets worden voorgelegd dient iedere deelset aan een representatieve steekproef voorgelegd te worden.
- Iedere ankeropgave wordt aan tenminste 1.000 leerlingen voorgelegd.
- De ankeropgaven zijn zodanig over de toets verspreid dat de positie minimale invloed heeft op de prestatie van de leerlingen op deze opgaven. Concreet betekent dit dat de ankeropgaven niet helemaal voorin of achterin de toets geplaatst zijn op een bepaalde afnamedag of toetstaak³.
- De ankeropgaven vallen qua inhoud en vormgeving niet herkenbaar op tussen de reguliere opgaven in de toets. Ze worden niet in een afwijkend blok afgenomen maar vermengd met reguliere opgaven van hetzelfde vaardigheidsdomein.
- Toetsaanbieders kunnen minimale aanpassingen doen in de ankeropgaven om deze afneembaar te maken voor hun eigen doorstroomtoets en ze beter te integreren, mits goedgekeurd door het CvTE. Dit kan zijn:
 - Aanpassingen in lay-out zodat de opgave past in de toetsvorm of de afnamespeler van toetsaanbieder.
 - Kleine tekstuele wijzingen in tekst ter bevordering van consistentie over de toets.
 - Kleine wijzigingen in de opmaak van afbeeldingen ten behoeve van de consistentie.Deze lijst kan verder uitgebreid en geconcretiseerd worden. Ook kan deze lijst aangevuld worden met aanpassingen die *niet* doorgevoerd mogen worden, omdat is gebleken dat een dergelijke aanpassing tot DIF kan leiden. Dit zal gaandeweg gebeuren binnen de werkgroep gezamenlijk anker. De aanpassingen mogen geen itembias tot gevolg hebben. De gemaakte afspraken worden vastgelegd in een overzicht werkafspraken met de toetsaanbieders.
- De gezamenlijk ankeropgaven tellen mee in de vaardigheidsschattingen en totaalscores van de leerlingen⁴, tenzij de opgaven na analyse afgekeurd zijn bij de kalibratie t.b.v. de landelijke normering zoals omschreven in 2.2.1E.

2.3.5. Geheimhoudingsverplichting

Het is niet toegestaan vastgestelde ankeropgaven zonder overleg met CvTE elders te gebruiken. Toetsaanbieders zullen aan derden geen informatie, kennis of gegevens openbaren of doen openbaren, op welke wijze dan ook, met betrekking tot de ankerbank, het gezamenlijk anker en alle andere informatie die aan hen bekend is geworden en waarvan zij redelijkerwijs moeten begrijpen dat deze informatie van vertrouwelijke aard is.

Deze geheimhoudingsplicht blijft van kracht op het moment dat een toetsaanbieder niet langer een doorstroomtoets aanbiedt.

³ Zie ook toelichting T2 in Regeling beoordelingskader doorstroomtoets PO.

⁴ Zie ook hoofdstuk 3 uit bijlage 1 uit de Regeling beoordelingsnormen doorstroomtoetsen PO.



2.4. Toetsmatrijs

Tijdens de selectie van de jaarsets van het gezamenlijk anker worden de volgende uitgangspunten gehanteerd:

- Iedere jaarset van het gezamenlijk anker bevat voor ieder verplicht onderdeel 20 opgaven.
- Binnen ieder domein bestaat de jaarset uit een gelijke verhouding opgaven van referentieniveau 1F en 2F (voor Taalverzorging en Lezen) /1S (voor Rekenen).

De toetsmatrijs geldt als *richtlijn* voor de vaststelling van iedere jaarset, hier kan van jaar tot jaar in beperkte mate van afgeweken worden, bijvoorbeeld als de vulling van de bank niet toereikend is:

Lezen

Taakuitvoering	Aantal
Begrijpen	6
Evalueren	2
Interpreteren	6
Opzoeken	3
Samenvatten	3

Er zijn minstens twee zakelijke teksten en één fictionele/narratieve/literaire tekst. Van de zakelijke teksten is er minstens één informatief en één instructief of betogend.

Taalverzorging

Onderdeel	Categorie	Aantal
Interpunctie (5)	Leestekens	5
Spelling niet-werkwoorden (8)	Lettergreepgrenzen	2
	Morfologisch	4
	Overige regels ¹	2
Spelling werkwoorden (7)	Morfologisch	7

¹ Onder overige regels worden spelambigue woorden (logografische spelling), schrijfwijze van tussenklanken en gebruik van trema en koppelteken bedoeld (Meijerink, 2009).

Rekenen

Domein	Onderdeel	Aantal
Getallen (7, >1 kaal, >2 context)	Notatie, taal en betekenis	1
	Verbanden	2
	Gebruiken	4
Verhoudingen (5, >1 kaal, >2 context)	Verbanden	2
	Gebruiken	3
Meten en meetkunde (5)	Verbanden	3
	Gebruiken	2
Verbanden (3)	Verbanden	1
	Gebruiken	2

3. Landelijke (toetsoverstijgende) normering

Na afname van de varianten van de doorstroomtoetsen van het vigerende kalenderjaar, start de landelijke (toetsoverstijgende) normering. In deze Regeling Beoordelingsnormen Doorstroomtoetsen po worden normeringsstappen beschreven. Het CvTE stelt vervolgens voor ieder van de afgenomen doorstroomtoetsen de normering vast.

Om recht te doen aan een betrouwbare en valide inschatting van het vaardigheidsniveau van leerlingen, het vergroten van de vergelijkbaarheid, in het geval van een calamiteit of anderszins gewichtige reden, kan het CvTE besluiten om van de in deze Regeling Beoordelingsnormen Doorstroomtoetsen po beschreven wijze van normering af te wijken.

In het geval van ingrijpen door het CvTE omdat de normering bij een specifieke doorstroomtoets niet kan worden uitgevoerd, kan dat consequenties hebben voor de erkenning of toelating van de betreffende doorstroomtoets voor het erop volgende jaar.

3.1. Planning en aanleveren en data

In onderstaande tabel staat de planning van de normeringsfasen beschreven. Een gedetailleerde beschrijving van de werkzaamheden is beschreven in de volgende paragrafen. Het is uiteraard mogelijk dat een stap eerder is afgerond dan de deadline. Het CvTE kan in bijzondere gevallen genooddakt zijn de genoemde data aan te passen en zal in dat geval de nieuwe planning met de toetsaanbieders delen.

Fase	Code	Stap	Wie	Deadline
1. Datalevering				
		Leveren benodigde gegevens	Toetsaanbieders plaatsen data op de terminal-server	Maandag na de laatste week van de officiële afnameperiode (09:00)
2. Kalibratie gezamenlijk ankeropgaven				
		Levering resultaten, na akkoord CvTE	de adviseur plaatst resultaten op de terminal-server	donderdag na de laatste week van de officiële afnameperiode (09:00)
Re 1. (optioneel) Update datalevering				
		Definitieve datalevering toetsaanbieders (optioneel)	Toetsaanbieders plaatsen data op de terminal-server	vrijdag na de laatste week van de officiële afnameperiode (09:00)
3. Normeringsprocedure IRT				
		Voorlopige normeringsvergadering CvTE	de adviseur aan CvTE	Dinsdag, de week na datalevering (09:00)
4. Definitieve beslissing normering				
		CvTE beslist over de definitieve normering, vaststelling cesuren toetsadviezen en referentieniveaus	de adviseur en CvTE	Dinsdag, de week na datalevering (uiterlijk 16:00)
		Presentatie van definitieve normering aan afstemmingsoverleg	de adviseur aan Toetsaanbieders	Dinsdag, de week na datalevering (16:00-18:00)

3.2. Datalevering

De toetsaanbieder is verantwoordelijk voor het correct, tijdig, volledig en veilig aan het CvTE beschikbaar stellen van consistente data uit de operationele afname van de varianten die zijn afgenomen in het betreffende jaar.

De toetsaanbieder levert voor de landelijke normering de voorgeschreven gegevens over de afgenomen doorstroomtoets op de terminal-server van het CvTE. Iedere toetsaanbieder heeft een map campagne van het betreffende jaar waarin de data geplaatst wordt.

De data (1.) dient op de in de tabel opgenomen datum aangeleverd te zijn. Pas wanneer de data gecontroleerd is, kan de kalibratie van de gezamenlijk ankeropgaven gestart worden. Daarna wordt eventueel nieuwe, aangevulde data (Re 1.) aangeleverd. Voorafgaand aan de campagne zal aan iedere toetsaanbieder gevraagd worden of zij de intentie hebben van deze gelegenheid gebruik te maken. Het format en de controle van deze optionele levering zijn gelijk aan die bij de initiële aanlevering.

Wanneer de controles van de data interne inconsistenties uitwijzen, wordt de toetsaanbieder hiervan op de hoogte gesteld. Het CvTE en de adviseur zullen de data niet bewerken anders dan voorzien en beschreven is in deze regeling. Wanneer er vertragingen in aanlevering van de data optreden kan het normeringsproces vertraging oplopen.

De toetsaanbieder levert een zo compleet mogelijke dataset, die bestaat uit ruwe scores en andere kenmerken van leerlingen van zowel het regulier basisonderwijs als speciaal (basis)onderwijs, maar exclusief NOB-scholen en particuliere scholen. Minimaal moet de dataset bij de uiteindelijke datalevering resultaten van 50% van het totale aantal inschrijvingen van een doorstroomtoets bevatten. Alle dataleveringen bevatten minstens 1.000 observaties uit het regulier basisonderwijs per gezamenlijk ankeropgave uit de verplicht op te nemen jaarset van het gezamenlijk anker. Leerlingen die speciale toetsversies maken, zoals bijvoorbeeld bij de toets voor blinden of slechthorenden of gebruik maken van de rekenkaart, worden niet geleverd en tellen niet mee in de hiervoor genoemde eisen. Leerlingen die een reguliere variant maken maar bijvoorbeeld van een verklanking gebruikmaken dienen wel

geleverd te worden. Afgebroken en ongeldige afnames vormen geen deel van de datalevering door de toetsaanbieder.

Voordat de ruwe scores worden geleverd door de toetsaanbieder, heeft de toetsaanbieder een itemanalyse uitgevoerd om eventuele sleutfouten en -uitbreidingen te signaleren. In de dataset die op de terminal-server wordt aangeleverd zijn deze fouten reeds gecorrigeerd door middel van sleutelwijzigingen of neutralisaties. Als de populatie van een toetsaanbieder een voor de itemkalibratie significant aandeel s(b)o leerlingen bevat, voert de toetsaanbieder een DIF-analyse uit op onderwijstype (bo versus s(b)o) en voert eventuele ingrepen om deze te remediëren (bijvoorbeeld door middel van herlabeling) uit alvorens de data aan te leveren.

De gegevens die geleverd zijn ten behoeve van de landelijke normering kunnen ook gebruikt worden voor nadere analyses door CvTE of de adviseur als daar behoefte aan is.

Indien de aan het CvTE tijdens de normeringscampagne geleverde data nog niet het volledige leerlingaantal bevat levert de aanbieder uiterlijk de eerste maandag van de landelijk vastgestelde rapportageperiode de volledige data aan.

3.2.1. Data en structuur

De toetsaanbieder levert voor de normering:

- Eén bestand met leerlinggegevens.
- Voor ieder onderdeel een apart bestand met scoringsgegevens.
- Eén bestand met p-waarden en aantal observaties van de items die opgenomen zijn in de analysedata (ter controle). Itemcodes dienen uniek te zijn over onderdelen heen.
- Representativiteitsgegevens in de vorm van alle bekende schooladviezen.

Optioneel: een lijst met items die niet meetellen in de scoring.

De data worden aangeleverd in csv (uitgaande van de Nederlandse situatie, dus puntkomma-gescheiden en komma als decimaalscheidingsteken). De teruggeleverde databestanden gebruiken hetzelfde format. De toetsaanbieders leveren geen aparte toetsbeschrijving.

3.2.2. Leerlinggegevens

In het bestand met leerlinggegevens dienen de volgende variabelen opgenomen te worden:

Kolomnaam	Omschrijving
person_id	Leerlingidentificatiecode, numeriek of string. Niet herleidbaar naar persoonsgegevens.
populatie	Naam van populatie/toetsboekje/covariaat, optioneel en alleen van toepassing als er daadwerkelijk populaties zijn die aanleiding geven tot verschillende groepering in de MML-schatting.
schooltype	Onderwijstype, numeriek: bo = 1 sbo = 2 so = 3 overig = 4
schoolcode	RIO instellingscode plus vestigingscode, in formaat 99XX00. Voorheen bekend als BRIN-nummer plus vestigingscode. Bij leerlingen van scholen waarvan in totaal minder dan 5 leerlingen zijn ingeschreven bij de toetsaanbieder mag waarde "<5" gebruikt worden.
schooladvies	Voorlopig schooladvies, numeriek: pro = 1 vmbo bb = 2 vmbo bb/kb = 3 vmbo kb = 4 vmbo kb/gt = 5 vmbo gt = 6 vmbo gt/havo = 7 havo = 8 havo/vwo = 9 vwo = 10 onbekend = 0 of >10 Eventuele toevoeging aan het voorlopig schooladvies van leerwegondersteunend onderwijs (lwoo) mag genegeerd worden bij de indeling. Een voorlopig schooladvies van vmbo gt met lwoo wordt dan dus gewoon vmbo gt. Waarden die buiten bovengenoemde categorisering vallen (bijvoorbeeld vso) mogen als onbekend opgenomen worden.

Als naamgeving van dit bestand wordt verwacht:
TOETS_leerlingen.csv

3.2.3. Scoringsgegevens

Voor de scoringsgegevens worden zowel wide-format als long-format data verwerkt. Indien data in wide-format worden aangeleverd dienen per onderdeel de volgende scoringsgegevens opgenomen te worden:

Kolomnaam	Omschrijving
person_id	Leerlingidentificatiecode, numeriek of string
item_ids...	Voor ieder item een kolom met als naam het label van het item, met daarin de score: 0 = incorrect (inclusief overgeslagen of leeggelaten items) 1 = correct leeg = <i>missing by design</i>

In wide-format data staat het complete resultaat van een leerling op één regel. Indien data in long-format worden aangeleverd dienen per onderdeel de volgende scoringsgegevens opgenomen te worden:

Kolomnaam	Omschrijving
person_id	Leerlingidentificatiecode, numeriek of string
item_id	Itemlabel
item_score	Item score: 0 = incorrect (inclusief overgeslagen of leeggelaten items) 1 = correct <i>missing by design</i> wordt niet geleverd

In long-format data vormt ieder leerling-item-paar een aparte regel.

Als naamgeving van het bestand met scoringsgegevens wordt verwacht:
TOETS_[naam onderdeel in hoofdletters].csv

Voor de verplichte onderdelen worden de namen LEZEN, REKENEN en TAAL verwacht.

3.2.4. Representativiteitsgegevens

Naast een controle op de representativiteit van de geleverde dataset, worden deze gegevens gebruikt om de landelijke prestaties op de doorstroomtoetsen te vergelijken met voorgaande jaren. De totale leerlingaantallen worden gevalideerd met behulp van inschrijfggegevens van DUO. In het bestand met representativiteitsgegevens dienen de volgende variabelen opgenomen te worden:

Kolomnaam	Omschrijving
schooladvies	Voorlopig schooladvies, zelfde indeling als in bestand met leerlinggegevens
n	Aantal leerlingen met dit schooladvies in de <i>complete</i> dataset, ook van leerlingen die nog niet gescoord zijn, voor zover bekend en van alle onderwijstypen samen

Als naamgeving van het bestand met representativiteitsgegevens wordt verwacht:
TOETS_representativiteit.csv

3.2.5. Controlegegevens

In het controlebestand met p-waarden dienen tot slot de volgende gegevens te worden aangeleverd:

Kolomnaam	Omschrijving
item_id	Itemlabel
onderdeel	Naam van het onderdeel in hoofdletters, overeenkomend met de onderdeelnamen in de namen van de scoringsgegevens
n	Aantal observaties (alle onderwijstypen samen)
p_waarde	p-waarde van het item (alle onderwijstypen samen)

Als bestandsnaam voor het controlebestand met p-waarden wordt verwacht:
TOETS_controle.csv

3.2.6. Optioneel: lijst van opgaven die niet meetellen

Als er opgaven zijn die niet meetellen in de scoring van de toets, zoals zaai-opgaven, kunnen deze meegeleverd worden in een opgavenlijst. Deze opgaven kunnen mogelijk als extra ankeringsvragen dienen tussen afnamejaren. Ook kunnen gezaaide opgaven ten behoeve van de samenstelling van het



gezamenlijk anker geëvalueerd worden. Deze opgaven zullen niet gebruik worden in de kalibratie en normering ten behoeve van de doorrekening door de adviseur (stap 2 in het normeringsproces). De opgaven worden aangeleverd in een plat tekstbestand waarbij ieder label op een nieuwe regel staat. Als er geen lijst wordt aangeleverd, wordt er vanuit gegaan dat alle opgaven meetellen voor de scoring.

Als naamgeving van het bestand met no score-items wordt verwacht:
TOETS_noscore.dat

3.2.7. Datacontroles

de adviseur voert de volgende controles uit op de aangeleverde data:

- Zijn alle bestanden aanwezig met de juiste namen en kolommen?
- Zijn leerlingidentificatiecodes uniek?
- Zijn alle onderwijstypes toegestaan?
- Zijn opgaven per leerling uniek?
- Kloppen de p-waarden en aantallen observaties met het controlebestand?
- Bevat ieder verplicht onderdeel ankeropgaven? (labels conform parameterbestand)
- Is het design verbonden?
- Check op correcte leerlingkoppeling door correlatiecoëfficiënten van verschillende onderdelen uit te rekenen. Bij een correlatie lager dan 0,4 koppelt de adviseur terug aan de toetsaanbieder dat er een lage correlatie is geconstateerd, met het verzoek de koppeling op leerlingniveau te controleren. Wanneer de toetsaanbieder aangeeft dat de lage correlatie correct is zal de adviseur de normering uitvoeren op de aangeleverde data.

3.3. Normeringsprocedure

Nadat de gescoorde data door de adviseur zijn gecontroleerd en geanalyseerd volgt de landelijke (toetsoverstijgende) normeringsprocedure.

De normering wordt door de adviseur uitgevoerd met behulp van het gezamenlijke anker. De berekeningen worden uitgevoerd in daarvoor voorgeschreven gevalideerde software. De normeringsprocedure bestaat uit de volgende stappen:

1. kalibratie gezamenlijk ankeropgaven
2. doorrekening normering
3. toepassing normering

Periodiek worden de onderdeelgewichten en de cesuren voor de referentieniveaus en de toetsadvies-categorieën geanalyseerd en zo nodig geactualiseerd. Hiervoor worden onder andere gegevens uit het meest recente toelatings- en doorstroomonderzoek gebruikt.

3.3.1. Stap 1: kalibratie gezamenlijk ankeropgaven

Voor ieder van de wettelijk verplichte onderdelen Lezen, Rekenen en Taalverzorging wordt de afnamedata van alle toetsaanbieders samengenomen. Deze data worden gekalibreerd in het 2PL-itemresponsmodel, waarbij iedere toets als afzonderlijke populatie in de MML⁵-kalibratie wordt behandeld. Alleen leerlingen van regulier basisonderwijs worden hierin meegenomen. Verder worden opgaven niet gebruikt wanneer:

- de opgave minder dan 200 observaties heeft;
- er sprake is van een item met slechts één responscategorie
- de opgave een discriminatieparameter in een initiële kalibratie van het 2PL-model heeft van $a < 0.1$ en $a > 10$.

In deze kalibratie worden gegevens uit alle afnamejaren van de eind- en doorstroomtoetsen vanaf 2018 gebruikt om een koppeling te maken naar de referentiesets.

Alle opgaven die een koppeling vormen tussen eind- en doorstroomtoetsen in de verschillende afnamejaren worden gecontroleerd op mogelijk differentieel functioneren (DIF) tussen toetsen onderling. Dit gebeurt door middel van een statistische analyse naar hand van modelpassingsmaten en een visuele inspectie van de responscurves. Bij het constateren van differentieel functioneren wordt hiervoor gecompenseerd met de ingrepen zoals besproken in 2.2.1E. De bevindingen worden teruggeraapporteerd aan de toetsaanbieders.

⁵ Marginal maximum likelihood (MML) is een schattingsmethode om tot itemparameters te komen in een kalibratie.

Cesuurpunten referentieniveaus

De cesuurpunten voor de referentieniveaus op de vaardigheidsschaal die verkregen is in de hierboven genoemde kalibratie worden gebaseerd op een equipercntiel-equivalering met de gerealiseerde behaalde percentages referentieniveaus van de Centrale Eindtoets (papier en digitaal) van 2018

Cesuurpunten toetsadviezen

De toetsadviezen worden gebaseerd op de gemiddelde latente vaardigheid (glv). Dit is een samenvoeging van de vaardigheden van de onderdelen die meetellen voor het toetsadvies. De wegingsfactoren van de verschillende onderdelen in de glv zijn:

lezen	rekenen	taalverzorging
0,30	0,45	0,25

De cesuurpunten voor de toetsadviescategorieën op de glv-schaal worden gebaseerd op een equipercntiel-equivalering met de gerealiseerde percentages toetsadviezen van de papieren Centrale Eindtoets van 2019, waarbij rekening gehouden wordt met een verschuiving van de grenswaarde tussen pro/vmbo bb en vmbo bb/kb die in 2021 heeft plaatsgevonden.

Na het uitvoeren van de kalibratieprocedure worden de parameters van de gezamenlijk ankeropgaven en de cesuren voor de toetsadviezen en referentieniveaus geaccordeerd door het CvTE en gedeeld met de toetsaanbieders.

Teruglevering itemparameters gezamenlijk anker

de adviseur levert de toetsaanbieders een csv-bestand met gekalibreerde itemparameters, zowel van de bestaande als de nieuwe opgaven van het gezamenlijk anker. Hierin staan de volgende kolommen:

Kolomnaam	Omschrijving
item_id	Itemlabel
item_score	Score van het item (1)
beta	Moeilijkheidsparameter
SE_beta	Standaardfout op moeilijkheidsparameter
alpha	Discriminatieparameter
SE_alpha	Standaardfout op discriminatieparameter
onderdeel	Het onderdeel van het item

Als bestandsnaam wordt gebruikt:
ankerparameters_JAAR.csv

Teruglevering ingrepen kalibratie gezamenlijk anker

de adviseur levert een csv-bestand met een overzicht van alle ingrepen die zijn uitgevoerd in de opgaven die voor een koppeling zorgen tussen de verschillende eind- en doorstroomtoetsen in alle afnamejaren ten behoeve van de kalibratie van het gezamenlijk anker.

Kolomnaam	Omschrijving
jaar	Het afnamejaar waarin de ingreep wordt toegepast
toets	De eind- of doorstroomtoets waarin de ingreep wordt toegepast
item_id	Itemlabel
onderdeel	Het onderdeel behorende bij het item
actie	De ingreep die wordt uitgevoerd, kan waarden "los" (loskoppelen door middel van omlabeling) en "uit" hebben. Zie ook 2.2.1E
opmerking	Eventuele aanvullende onderbouwing

Als bestandsnaam wordt gebruikt:
dif_ingrepen_JAAR.csv

Teruglevering uitgeschakelde items kalibratie gezamenlijk anker

Tevens levert de adviseur dan aan de toetsaanbieders een bestand met uitgeschakelde items die een te lage of te hoge discriminatieparameter kregen in de initiële 2PL-kalibratie.

Kolomnaam	Omschrijving
item_id	Itemlabel
onderdeel	Het onderdeel behorende bij het item
alpha	Discriminatieparameter
beta	Moeilijkheidsparameter
SE_alpha	Standaardfout op discriminatieparameter
SE_beta	Standaardfout op moeilijkheidsparameter

Als bestandsnaam wordt gebruikt:
TOETS_uitgeschakeld_ankerkalibratie.csv

3.3.2. Stap 2: doorrekening normering doorstroomtoetsen

De toetseigen opgaven van de verplichte onderdelen van iedere toetsaanbieder worden apart gekalibreerd in het 2PL-itemresponsmodel, waarbij het gezamenlijk anker wordt gefixeerd op de in stap 1 (kalibratie anker) verkregen parameters. Hierdoor worden de opgaven van iedere toets op dezelfde, vergelijkbare schaal gebracht. In deze procedure worden dezelfde eisen gesteld aan opgaven als in stap 1. In deze kalibratie worden leerlingen uit alle onderwijstypen meegenomen. De leerlingen uit de verschillende onderwijstypen worden als verschillende groepen behandeld in de MML-schatter. Met de parameters voor alle opgaven in de toets worden vervolgens vaardigheidsschattingen uitgevoerd met behulp van Warm's gewogen waarschijnlijkheidsschatter (wml) voor alle leerlingen voor ieder verplicht toetsonderdeel. De vaardigheidsschattingen worden vergeleken met de in stap 1 berekende cesuurpunten voor de referentieniveaus 1F en 2F/1S voor de drie wettelijk verplichte onderdelen. Voor iedere leerling van wie de gegevens zijn aangeleverd door de toetsaanbieder worden de vaardigheden op alle onderdelen die meetellen voor het toetsadvies samengevoegd tot de glv. De glv's worden vervolgens vergeleken met de in stap 1 berekende cesuurpunten voor de toetsadviescategorieën om op basis hiervan een toetsadvies toe te kennen.

Teruglevering uitgeschakelde items

Dit bestand bevat de items die uitgeschakeld zijn in de IRT-kalibratie omdat ze een discriminatieparameter met een waarde lager dan 0.1 of hoger dan 10 hebben gekregen in de initiële 2PL-kalibratie of te weinig observaties hadden. Als er geen items uitgeschakeld zijn wordt een leeg bestand opgeleverd. De volgende kolommen zijn opgenomen:

Kolomnaam	Omschrijving
item_id	Itemlabel
onderdeel	Het onderdeel behorende bij het item
alpha	Discriminatieparameter
beta	Moeilijkheidsparameter
SE_alpha	Standaardfout op discriminatieparameter
SE_beta	Standaardfout op moeilijkheidsparameter

Als bestandsnaam wordt gebruikt:
TOETS_uitgeschakeld.csv

Teruglevering itemparameters

Dit bestand bevat de itemparameters zoals geschat in de IRT-normering volgens het 2PL-model. De volgende kolommen zijn opgenomen:

Kolomnaam	Omschrijving
item_id	Itemlabel
onderdeel	Het onderdeel behorende bij het item
alpha	Discriminatieparameter
beta	Moeilijkheidsparameter
SE_alpha	Standaardfout op discriminatieparameter
SE_beta	Standaardfout op moeilijkheidsparameter

Als bestandsnaam wordt gebruikt:
TOETS_itemparameters.csv

Teruglevering populatieparameters

Bevat de MML-populatieparameters zoals geschat in de IRT-normering in het 2PL-model voor alle onderdelen. De volgende kolommen zijn opgenomen:

Kolomnaam	Omschrijving
populatie	Naam van populatie/toetsboekje/covariaat indien van toepassing, anders de naam van de toets
group_n	Aantal leerlingen in de populatie
onderdeel	Het onderdeel
mu	Populatiegemiddelde
sd	Standaarddeviatie van de populatie
SE_mu	Standaardfout op het populatiegemiddelde
SE_sd	Standaardfout op de standaarddeviatie van de populatie

Als bestandsnaam wordt gebruikt:
TOETS_populatieparameters.csv

Teruglevering vaardigheidsschattingen

Bevat voor iedere leerling de vaardigheid op ieder onderdeel in het 2PL-model geschat via de wml-methode. Tevens zijn de berekende GLV's en de via de cesuren daaraan gerelateerde toetsadviezen hieraan toegevoegd. De volgende kolommen zijn opgenomen:

Kolomnaam	Omschrijving
person_id	Leerlingidentificatiecode
schooltype	Onderwijstypen, zelfde indeling als in bestand met leerlinggegevens
theta_[onderdeel]	Vaardigheidsschatting voor ieder onderdeel
glv	glv
toetsadvies	Toetsadvies volgens de IRT-normering

Als bestandsnaam wordt gebruikt:
TOETS_vaardigheden.csv

Teruglevering referentieniveaus

Bevat een overzicht van de behaalde referentieniveaus in de IRT-normering in het 2PL-model, waarbij alleen leerlingen uit regulier BO zijn meegenomen. De volgende kolommen zijn opgenomen:

Kolomnaam	Omschrijving
onderdeel	Het onderdeel
niveau	Het referentieniveau
perc_behaald	Percentage behaald

Als bestandsnaam wordt gebruikt:
TOETS_referentieniveaus.csv

Teruglevering toetsadviezen

Bevat een overzicht met behaalde toetsadviezen in de IRT-normering. Alleen leerlingen uit regulier BO worden meegenomen. De volgende kolommen zijn opgenomen:

Kolomnaam	Omschrijving
toetsadvies	Het toetsadvies
perc_behaald	Percentage behaald

Als bestandsnaam wordt gebruikt:
TOETS_toetsadviezen.csv

Controle

Met behulp van een populatieschatting op basis van de voorlopige schooladviezen wordt een vergelijking tussen de toetsaanbieders gemaakt van de prestaties op de referentieniveaus en de verdeling van toetsadviezen. Hiervoor zal conditioneel op voorlopig schooladvies een trekking worden

uitgevoerd uit de behaalde referentieniveaus en toetsadviezen van de volledige landelijke populatie (voor zover de schooladviezen bekend zijn).

3.3.3. Stap 3: toepassing normering doorstroomtoetsen

De in stap 2 (doorrekening normering doorstroomtoetsen) doorgerekende resultaten op de normeringssteekproef dienen slechts ter controle. Iedere toetsaanbieder past de vastgestelde normering, in de vorm van de in stap 1 (kalibratie gezamenlijk anker) vastgestelde parameterfixaties voor het gezamenlijk anker en vastgestelde cesuren voor toetsadviezen en referentieniveaus, zelf toe voor iedere deelnemende leerling ten behoeve van de rapportages. Indien een toets gebruikmaakt van een opgavebank of de steekproef te beperkt was, kunnen hierbij ook opgaven meegenomen worden die in de operationele afname van de doorstroomtoets minder dan 200 observaties hadden, mits voor die opgaven voldaan wordt aan de eisen voor itemparameterschattingen zoals genoemd in 5.2.1 doel B in het Beoordelingskader.

Nadat de toetsadviezen en referentieniveaus zijn uitgerekend voor de leerlingen in de met het CvTE en de adviseur gedeelde normeringssteekproef wordt de normering, inclusief de cesuren voor de referentieniveaus en de toetsadviescategorieën en de onderdeelgewichten, door het CvTE definitief vastgesteld.

3.3.4. Optionele onderdelen

Naast de wettelijk verplichte onderdelen worden ook de optionele onderdelen die meetellen in het toetsadvies van iedere toetsaanbieder door de adviseur gekalibreerd door middel van een voorgescreven kalibratiewijze met bijbehorende schattingsmethode, uit te voeren in de extern gevalideerde software, waarbij per toetsaanbieder dezelfde populaties als bij de wettelijk verplichte onderdelen worden gebruikt.

4. Normering bij inzet OCW Doorstroomtoets

Omdat een mogelijk inzetscenario van de OCW Doorstroomtoets het uitlekken van gezamenlijk ankeropgaven is, bevat de OCW Doorstroomtoets geen gezamenlijk ankeropgaven. In plaats van een parameterfixatie op het gezamenlijk anker, worden de vaardigheidsschalen in de normering van de OCW Doorstroomtoets gefixeerd door middel van een koppeling naar de Centrale Eindtoets en/of DOE. De OCW Doorstroomtoets bestaat namelijk uit opgaven die via pretestafnames zijn gekoppeld aan afnames van de Centrale Eindtoets. Omdat de Centrale Eindtoets wel het gezamenlijk anker bevatte, kan via deze route de OCW Doorstroomtoets op dezelfde schaal gebracht worden als het gezamenlijk anker. Hiertoe zal een kalibratie met het 2PL-model gebruikt worden. Als deze schalen zijn overgezet, kunnen de cesuurpunten voor de referentieniveaus en toetsadviezen zoals beschreven in deze regeling toegepast worden om tot de waarden voor leerlingrapportages te komen.

Ruwweg worden twee scenario's voor de inzet van de OCW Doorstroomtoets voorzien:

1. Een kleinschalige inzet vanwege een gelokaliseerde calamiteit, minder dan 1.000 leerlingen.
2. Een grootschalige inzet met hoge leerlingaantallen.

Bij een kleinschalige inzet zijn er te weinig observaties van de opgaven uit de OCW Doorstroomtoets voor een betrouwbare kalibratie. In dat geval zal de OCW Doorstroomtoets vooraf gekalibreerd worden op basis van pretestgegevens, en worden de normeringen voor de (leerling)rapportages op deze kalibraties gebaseerd. Bij een grootschalige inzet zal wel een nieuwe kalibratie uitgevoerd worden op basis van de operationele afname. Voorwaarde hiervoor is echter wel dat er voldoende tijd is voor de rapportagedeadline van 15 maart. Of dit haalbaar is wordt beoordeeld en besloten in het calamiteitenteam

Omdat het gezamenlijk anker niet wordt afgenomen in de OCW Doorstroomtoets, zal de toets niet gecontroleerd worden in de reguliere landelijke toetsoverstijgende normering. Indien schooladviezen beschikbaar zijn van de deelnemende leerlingen kunnen de resultaten van de OCW Doorstroomtoets wel meegenomen worden om inzicht te krijgen in de landelijke relatie tussen schooladviezen en toetsadviezen. Hiertoe levert Stichting Cito team DOE/ OCW Doorstroomtoets een csv-bestand met voor iedere leerling het onderwijstype, het schooladvies, en de geschatte latente vaardigheden voor ieder onderdeel (zie voor de dataomschrijving de tabellen in paragraaf 3.3 en 5.7).

5. Datalevering toelatings- en doorstroomonderzoek

5.1. Achtergrond

Voor de normering moet o.a. toelatings- en doorstroomonderzoek (hierna: doorstroomonderzoek)

verricht worden. Dit houdt in dat gekeken wordt hoe de resultaten die leerlingen behalen op een doorstroomtoets in groep 8 zich verhouden tot de schoolsoort waarop zij zich de eerste drie jaar van het voortgezet onderwijs (vo) bevinden. De voorspellende waarde van de toetsresultaten wordt hiermee bepaald. Om dit te kunnen doen is een koppeling nodig tussen twee typen data: toetsresultaten van individuele leerlingen op de doorstroomtoets in groep 8 en (doorstroom)gegevens van de plaatsing van deze leerling in het eerste, tweede en derde jaar van het vo.

5.2. Proces verwerking persoonsgegevens

Het doorstroomonderzoek voert de adviseur uit in opdracht van het CvTE. De databronnen hiervoor zijn toetsresultaten op leerlingniveau van de toetsaanbieders, en plaatsingsgegevens van het vo uit het Register Onderwijsdeelnemers (ROD) van DUO. Om de koppeling tussen deze twee databronnen te kunnen maken, moeten de gegevens terug te voeren zijn op individuele leerlingen en moeten er dus persoonsgegevens verwerkt worden.

Om de uitwisseling van persoonsgegevens tot een minimum te beperken zal voor de koppeling per individuele leerling een pseudoniem gebruikt worden. Door middel van het toepassen van een reeds bestaand pseudonimiseringsalgoritme kunnen toetsaanbieders en DUO zelf dit pseudoniem genereren. Doordat DUO en de toetsaanbieders vier dezelfde gegevens van een individuele leerling in dit pseudonimiseringsalgoritme stoppen, komt hetzelfde pseudoniem voor een individuele leerling eruit bij de twee verschillende bronnen. Op basis van dit pseudoniem kan de adviseur de gegevens koppelen. Er is bij gebruik van dit algoritme geen sleutel waarmee het pseudoniem weer teruggezet kan worden door een ander dan degene die ook beschikt over de persoonsgegevens zelf. Alleen DUO en de individuele toetsaanbieders die de gegevens hebben geleverd kunnen het pseudoniem weer koppelen aan de persoonsgegevens waarmee het pseudoniem is gegenereerd. Op die manier beschikt de adviseur onder verantwoordelijkheid van het CvTE over zo min mogelijk gevoelige gegevens om deze wettelijke taak goed uit te kunnen voeren.

5.3. Wijze van pseudonimiseren

De pseudoniemen worden met het hashingalgoritme SHA-256 gegenereerd. Om de pseudoniemen later te kunnen matchen is het van belang dat alle partijen exact dezelfde set aan persoonsgegevens gebruiken om tot hetzelfde pseudoniem te komen. De invoer bestaat uit de instellingscode, achternaam, toetsscore en geboortedatum van de leerling. Indien bij een leerling één van deze gegevens mist, dient de data van deze leerling niet gepseudonimiseerd te worden en wordt deze data niet geleverd aan de adviseur. Ditzelfde is het geval bij leerlingen voor wie bovenstaande gegevens identiek zijn, bijvoorbeeld bij tweelingen op dezelfde school met dezelfde toetsscore. Data van leerlingen die een doorstroomtoets maken in het buitenland of op private scholen worden niet geleverd. Leerlingen van SO, SBO en VSO worden wel gepseudonimiseerd en geleverd aan de adviseur.

De specificatie voor de vier invoergegevens van het pseudoniem is als volgt:

Veld	#	Betekenis en waardenbereik
Schoolcode	1	RIO instellingscode. Waarde altijd tekst van 4 karakters (2 cijfers en 2 letters). Dit komt overeen met de eerste vier tekens van de BRIN-code.
Achternaam leerling	2	Achternaam van de leerling. Waarde is het significante deel van de achternaam, dus zonder voorvoegsel en zonder de scheidingsspatie volgend op het voorvoegsel [NEN 1888:2002]. De waarde is maximaal 70 karakters. Indien het significante deel van de achternaam uit meer delen bestaat, die gescheiden zijn door een koppelteken ("‑"), wordt geen spatie geplaatst voor en/of na dit koppelteken. Bijvoorbeeld: "Van der Berg" wordt "Berg", "O'Brien" blijft "O'Brien", "Al Khadouri" wordt "Khadouri", "El-Idrissi" blijft "El-Idrissi", "Da Silva-Pereira" wordt "Silva-Pereira", "van der Berg-van Laar" wordt "Berg-van Laar".
Geboortedatum leerling	3	Datum van geboorte van de leerling; waarde volgens formaat EEEE-MM-DD.

Veld	#	Betekenis en waardenbereik
Toetsscore leerling	4	De toetsscore van de doorstroomtoets is een geheel getal (verplicht volgens PvE ROD-po) met de volgende waardenbereiken:
		IEP 750...800
		Route 8 100...300
		AMN 301...500
		Dia 321...390
		Leerling in beeld 151...200
		DOE 200...400
		OCW doorstroomtoets 200...400

Deze gegevens worden gescheiden door komma's samengevoegd tot een gegevenssleutel, waarna het hashingsalgoritme wordt toegepast. Bijvoorbeeld:

Gegevenssleutel: "99CV,Test,2021-01-01,500"

Pseudoniem: "9fc18ccca231fa1951282b4a53ba6b71d3901122eff663c95758ac207cae8903"

De gegevenssleutel heeft als tekencodering UTF-8, bij gebruik van een andere codering kan een ander pseudoniem ontstaan (met name bij diakrieten).

5.4. Specificatie bestandlevering

Dataverlevering als csv volgens Nederlandse conventies:

- Scheidingsteken ";"
- Teksten zonder aanhalingstekens
- Decimaalscheidingsteken: komma
- Eerste regel bevat kolomnamen

Zowel de kolomnamen als de waarden zijn hoofdlettergevoelig.

5.4.1. Levering door DUO

Naamgeving "DUO-dso-" + [YYYYMMDD] + ".csv"

- [YYYYMMDD] = datum aanmaak bestand
- Selectie: alle leerlingen eindtoets/doorstroomtoets van afgelopen 3 kalenderjaren

kolomnaam	beschrijving
Pseudoniem	Zie beschrijving 5.3
Toetssort	Uit waardelijst "Toetssort"
Ontheffing	J/N
Toetsscore	
Kalenderjaar	Jaar van afname doorstroomtoets (YYYY)
Referentieniveau rekenen	Uit waardelijst "Referentieniveau rekenen"
Referentieniveau lezen	Uit waardelijst "Referentieniveau taal"
Referentieniveau taalverzorging	Uit waardelijst "Referentieniveau taal"
Verblijfjaar	
Leerjaar	
Niveau	Uit waardelijst "NiveauVO"
Voorlopig schooladvies	Uit waardelijst "Schooladvies"
Definitief schooladvies	Uit waardelijst "Schooladvies"
Type onderwijs	Uit waardelijst "Type onderwijs"

5.4.2. Levering door toetsaanbieders

Naamgeving [Toetssort]-dso-" + [YYYYMMDD] + ".csv"

- [Toetssort] = uit waardelijst "Toetssort"
- [YYYYMMDD] = datum aanmaak bestand



kolomnaam	beschrijving
Pseudoniem	Zie beschrijving 5.3
Toetssoort	Uit waardelijst "Toetssoort"
Afnamemodus	Uit waardelijst "Afnamemodus"
Toetsscore	Toetsscore van de doorstroomtoets, zie ook de tabel in 5.4
Kalenderjaar	Jaar van afname doorstroomtoets (YYYY)
Referentieniveau_rekenen	Uit waardelijst "Referentieniveau Rekenen"
Referentieniveau_lezen	Uit waardelijst "Referentieniveau Taal"
Referentieniveau_taalverzorging	Uit waardelijst "Referentieniveau Taal"
GLV	De GLV zoals bepaald door de toetsaanbieder (dus inclusief eventuele optionele onderdelen die meetellen voor het toetsadvies), op basis van gewogen gemiddelde van onderdeelvaardigheden. Afronden op 3 decimalen.
vaardigheid_rekenen	Door de toetsaanbieder geschatte latente vaardigheid op de onderdeel
vaardigheid_lezen	Rekenen, Lezen en Taalverzorging, op basis van een WML in het 2PL op de schaal van het gezamenlijk anker. Afronden op 3 decimalen.
vaardigheid_taalverzorging	
Toetsadvies	Uit waardelijst "Toetsadvies"
schoolcode	RIO instellingscode plus vestigingscode, in formaat 99XX00. Voorheen bekend als BRIN-nummer plus vestigingscode. Bij leerlingen van scholen waarvan in totaal minder dan 5 leerlingen zijn ingeschreven bij de toetsaanbieder mag waarde "<5" gebruikt worden.
Type onderwijs	Uit waardelijst "Type onderwijs"

Waardelijst

veld	waarden
Toetssoort	AMN, DIA, DOE, OCW DOORSTROOMTOETS_, ICE, LEERLING_IN_BEELD, ROUTE_8
Afnamemodus	digitaal, papier
Referentieniveau Rekenen	L1F, 1F, 1S
Referentieniveau Taal	L1F, 1F, 2F
NiveauVO	pro, vmbo bb, vmbo kb, vmbo gl, vmbo gl-tl, vmbo tl, havo, vwo
Toetsadvies	pro/vmbo bb, vmbo bb/vmbo kb, vmbo kb/vmbo gl-tl, vmbo gl-tl/havo, havo/vwo, vwo
Schooladvies	pro, vmbo bb, vmbo bb/vmbo kb, vmbo kb, vmbo kb/vmbo tl-tl, vmbo gl-tl, vmbo gl-tl/havo, havo, havo/vwo, vwo
Type onderwijs	BO, SBO, SO



TOELICHTING

De Regeling beoordelingsnormen doorstroomtoetsen PO bevat algemeen verbindende voorschriften voor aanbieders van een tot het stelsel toegelaten doorstroomtoets po. In de bijlage van de regeling staan de processtappen beschreven voor de voorbereiding op de landelijke (toetsoverstijgende) normering, waaronder begrepen de pretestprocedure en de toetssamenstelling. Een doorstroomtoets kan alleen worden erkend als de procedure om te komen tot de beoordelingsnormen van de doorstroomtoets wordt gevolgd.

Het bevoegd gezag is verplicht om in ieder geval gebruik te maken van door het College erkende doorstroomtoets voor de onderdelen Nederlandse taal en rekenen. Deze onderdelen moeten op grond van artikel 45b, vierde lid, van de Wet op het primair onderwijs, artikel 48c, vierde lid, van de Wet op de expertisecentra verplicht getoetst worden. Het College kan ook de onderdelen van de toets beoordelen waarvoor niet is voorgeschreven dat zij onderdeel uitmaken van een doorstroomtoets.

De bijlage van de regeling bestaat uit voorschriften voor de normering waaronder een gedetailleerde uitwerkingen voor de uitvoering van de leveringen van de gegevens en de voorschriften met betrekking tot het ontwikkelen, bewaren en gebruik van de ankeropgaven als bedoeld in artikel 4, onder d, van het Toetsbesluit PO.

De bijlage van de regeling wordt met deze regeling vervangen. De tekst is beperkt geactualiseerd en verbeterd. Ook is de toetsmatrijs voor het gezamenlijk anker aangepast, omdat vanaf 2027 de onderdelen techniek en woordenschat niet langer verplicht zijn maar wel aangeboden kunnen worden. De wijziging regelt dat de matrijs ook voor de selectie voor het anker van 2027 toereikend is. Verder is de regeling aangepast om verschillende methodes van de kalibratie van het anker uit te kunnen voeren tijdens de normering.

Tot slot is een uitbreiding van de te leveren gegevens ten behoeve van de normering en rapportages toegevoegd.

*Het College voor toetsen en examens,
de voorzitter,
J.H. van der Vegt*