



Regeling van het College voor toetsen en examens van 9 februari 2024, nummer CvTE-24.00436, houdende vaststelling van de beleidsregel voor de technische specificaties van de levering van de gegevens voor de beoordeling van de doorstroomtoetsen en de normering (Normeringshandboek 2024)

Het College voor toetsen en examens,

Gelet op de artikelen 4:81 van de Algemene wet bestuursrecht, 3a, eerste lid, onderdelen f en g, van de Wet College voor toetsen en examens en 2 van Regeling beoordelingskader doorstroomtoets PO;

Besluit:

Artikel 1

Het College voor toetsen en examens stelt de beleidsregels als bedoeld in artikel 2 van Regeling beoordelingskader doorstroomtoets PO en artikel 2 van de Regeling beoordelingsnormen doorstroomtoetsen PO vast als in de bijlage bij dit besluit.

Artikel 2

Deze regeling treedt in werking met ingang van de dag na de datum van uitgifte van de Staatscourant waarin zij wordt geplaatst en vervalt met ingang van 1 januari 2025.

Artikel 3

Deze beleidsregel wordt aangehaald als: Het normeringshandboek 2024.

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. HET NORMERINGSHANDBOEK VOOR DE DOORSTROOMTOETS PO 2024 (BIJLAGE ALS BEDOELD IN ARTIKEL 1)

HANDBOEK NORMERING 2024

1 INLEIDING

De Regeling beoordelingsnormen doorstroomtoetsen po bevat algemeen verbindende voorschriften voor aanbieders 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 technische specificaties voor de levering van de gegevens ten behoeve van de normering en operationele afspraken zijn een nadere uitwerking van de regeling beoordelingsnormen en worden beschreven in dit handboek normering.

In de afzonderlijke hoofdstukken van dit handboek worden operationele afspraken beschreven rondom de verschillende fasen van het normeringsproces. Het beschrijft procedures aangaande dataverwerking, analyses en hoe resultaten tot beslissingen leiden. Ook de rollen van de betrokkenen (zoals Stichting Cito, het CvTE) en de aanbieders van de doorstroomtoetsen in elk van de fasen komt aan bod. Door deze afspraken te expliciteren in het normeringshandboek ligt zoveel mogelijk van tevoren vast wat er van ieder verwacht wordt. Het eindresultaat van deze stappen is een definitieve normering.

De normering van de doorstroomtoets OCW wordt uitgevoerd door Stichting Cito team DOE/doorstroomtoets OCW, en valt buiten de procedures zoals opgesteld voor de landelijke normering in dit handboek. In hoofdstuk 6 wordt een korte toelichting gegeven op de normering van de doorstroomtoets OCW.

2 PLANNING

In onderstaande tabel staat de planning van de normeringsfasen beschreven. Een gedetailleerde beschrijving van de werkzaamheden is te vinden in het bijbehorende hoofdstuk. Bij de beschrijving wie betrokken is bij een fase worden afkortingen gebruikt (AA= Alle aanbieders; SC= Stichting Cito). Het is uiteraard mogelijk dat een stap eerder is afgerond dan de deadline. Het CvTE kan in bijzondere gevallen genoodzaakt 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				
	1.1	Leveren benodigde gegevens	AA aan SC	ma 19 februari (09:00)
	a	Data volgens afgesproken format		
	b	Bekende problemen met (anker)items melden		
	1.2	Data check: aanpassingen / meer informatie nodig?	SC	ma 19 februari (12:00)
		Per aanbieder: juiste / meer informatie aan leveren (kan vertragen)	Evt. aanbieder	ma 19 februari (17:00)
		Bij alle aanbieders de checks akkoord? Naar fase 2		
2. Kalibratie gezamenlijk ankeritems				
	2.1	Schaling van de (anker)opgaven	SC	
	a	Per referentieschaal: ankeropgaven evalueren	SC	wo 21 februari (09:00)
	2.2	Levering resultaten	SC aan AA	do 22 februari (09:00)
	a	Acties bij afwijkend functionerende items (globaal en per aanbieder)		
	b	Itemparameters gezamenlijk ankeritems		
Re 1. (optioneel) Update datalevering				
	Re1.1	Definitieve datalevering aanbieders (optioneel)	AA aan SC	vr 23 februari (09:00)
	Re1.2	Data check: aanpassingen / meer informatie nodig?	SC	vr 23 februari (10:00)
		Per aanbieder: juiste / meer informatie aan leveren (kan vertragen)	Evt. aanbieder	vr 23 februari (12:00)
		Bij alle aanbieders de checks akkoord? Naar fase 3		
3: Normeringsprocedure IRT				
	3.1	Berekening van de normering via IRT	SC	
	a	Normering referentieniveaus	SC	
	b	Normering toetsadviezen	SC	

Fase	Code	Stap	Wie	Deadline
	3.2	Voorgenomen normering delen met aanbieders (check op fouten) Toelichting advies	SC aan aanbieders	Ma 26 februari (13:00)
	3.3	Voorstel/presentatie voor normering aanleveren CvTE: Voorlopige normeringsvergadering	SC aan CvTE	di 27 februari (09:00)
5. Definitieve beslissing normering				
	5.1	CvTE beslist over de definitieve normering	SC CvTE	Di 27 februari (uiterlijk 16:00)
	5.2	Presentatie van definitieve normering aan afstemmingsoverleg	SC aan AA	Di 27 februari (16:00-18:00)

3 DATALEVERING

De data dient op **19 februari om 9:00** aangeleverd te zijn. Pas wanneer de data gecontroleerd is, kan fase 2 (kalibratie gezamenlijk ankeritems) gestart worden. Daarna wordt **optioneel op 23 februari voor 9:00** nieuwe, aangevulde data aangeleverd. Het is geen verplichting, maar voor de campagne zal aan iedere aanbieder 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. Het is de bedoeling dat zo veel mogelijk leerlingen van alle onderwijstypen worden geleverd, dus ook speciaal onderwijs

Wanneer de controles uitwijzen dat er een verschil is tussen wat gevonden wordt en wat verwacht wordt, zal de aanbieder zo spoedig mogelijk benaderd worden. Het is de verantwoordelijkheid van de aanbieder de data correct aan te bieden. Stichting Cito zal de data niet bewerken anders dan voorzien en beschreven is in de analyseprocedures. Wanneer er vertragingen in aanlevering van de data optreden kan het normeringsproces vertraging oplopen.

Het is de bedoeling dat een zo compleet mogelijke dataset geleverd wordt. Minimaal moet de dataset bij de tweede datalevering 50% van het totale aantal inschrijvingen van een toets bevatten, met minstens 1000 observaties per gezamenlijk ankeritem uit de jaarset. Leerlingen die speciale toetsversies maken, zoals bijvoorbeeld bij de toets voor blinden of slechthorenden, hoeven niet geleverd te worden en tellen niet mee in deze eisen. Leerlingen die bijvoorbeeld een verklanking van de toets maken dienen wel geleverd te worden.

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

De data zal aangeleverd worden via de terminal server van het CvTE. Iedere aanbieder heeft een map campagne_2024 waarin de data geplaatst wordt.

3.1 LANDELIJKE NORMERING

3.1.1 Data en structuur

De aanbieder levert voor de normering:

- Eén bestand met leerlinggegevens.
- Voor ieder onderdeel een apart bestand met toetsscores.
- 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 aanbieders leveren geen aparte toetsbeschrijving.

3.1.2 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.



Kolomnaam	Omschrijving
schooltype	Schooltype, numeriek: BO = 1 SBO = 2 SO = 3 overig = 4
schooladvies	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

Als naamgeving van dit bestand wordt verwacht:
TOETS_leerlingen.csv

3.1.3 Scoringsgegevens

Voor de scoringsgegevens wordt zowel wide-format als long-format data verwerkt. Indien data in wide-format wordt 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 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 wordt 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 1 = correct <i>missing by design</i> hoeft niet opgenomen te worden

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.

3.1.4 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 inschrijfgegevens van DUO. In het bestand met representativiteitsgegevens dienen de volgende variabelen opgenomen te worden:

Kolomnaam	Omschrijving
schooladvies	Schooladvies, zelfde indeling als leerling gegevens
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

3.1.5 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

3.1.6 Optioneel: lijst van no score-items¹

Als er items zijn die niet meetellen in de scoring van de toets kunnen deze meegeleverd worden in een itemlijst. Op deze manier is het mogelijk een kalibratie uit te voeren met behulp van alle (anker)items, maar kan de toets wel gescoord worden op dezelfde wijze als de aanbieder dat uitvoert. Deze items kunnen aangeleverd worden in een plat tekstbestand waarbij ieder itemlabel op een nieuwe regel staat. Als er geen lijst wordt aangeleverd, wordt er vanuit gegaan dat alle items meetellen voor de scoring.

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

TOETS_noscore.dat

3.1.7 Controles

Stichting Cito voert de volgende controles uit op de data en de kalibratie:

- Zijn alle bestanden aanwezig met de juiste namen en kolommen?
- Zijn leerlingidentificatiecodes uniek?
- Zijn alle schooltypes toegestaan?
- Zijn items per leerling uniek?
- Kloppen de p-waarden en aantallen observaties met het validatiebestand?
- Bevat ieder verplicht onderdeel ankeritems? (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 Stichting Cito terug aan de aanbieder dat er een lage correlatie is geconstateerd, met het verzoek de koppeling op leerlingniveau te controleren. Wanneer de aanbieder aangeeft dat de lage correlatie correct is zal Stichting Cito de normering uitvoeren op de aangeleverde data.
- DIF-analyse op gezamenlijk ankeritems door middel van een visuele inspectie en een impactanalyse rondom de cesuurpunten van de referentieniveaus. Wanneer een item DIF vertoont bij een aanbieder, wordt dit item bij deze aanbieder 'omgelabeld' en beschouwd als een regulier item, waarbij de itemparameters niet meer gefixeerd zullen zijn op de bekende itemparameters van het gezamenlijk anker. Deze procedure noemen we het 'loskoppelen' van het item.
- Met behulp van een populatieschatting op basis van de voorlopige schooladviezen wordt een vergelijking tussen de aanbieders gemaakt van de prestaties op de referentieniveaus en de verdeling van toetsadviezen. Hiervoor zal conditioneel op schooladvies een trekking worden uitgevoerd uit de behaalde referentieniveaus en toetsadviezen van de volledige landelijke populatie (voor zover de schooladviezen bekend zijn).

4 KALIBRATIE GEZAMENLIJK ANKERITEMS

De data van alle doorstroomtoetsaanbieders wordt samengenomen, waarbij iedere aanbieder als afzonderlijke populatie in de MML-kalibratie wordt meegenomen. Alleen leerlingen van regulier BO worden meegenomen.

In het najaar van 2021 heeft een herijking plaatsgevonden, waarbij de schalen en cesuurpunten van de

¹ In 2024 nog eenmalig mogelijk voor aanbieders waarmee coulance-afspraken hierover gemaakt zijn.

landelijke normering opnieuw bepaald zijn. De bestaande gezamenlijk ankeritems worden gefixeerd op deze schaal. De itemparameters op deze schaal worden voorafgaand aan de campagne gedeeld met de aanbieders. Met behulp van een kalibratie van het twee-parameter logistieke (2PL) model zullen de parameters van de nieuwe gezamenlijk ankeritems op de door de gefixeerde items gedefinieerde schaal geschat worden. De kalibratie wordt uitgevoerd met behulp van het R-package *dexterMML*².

Bij de kalibratie zijn er twee redenen om op voorhand opgaven uit te sluiten van de kalibratie:

- items met minder dan 200 observaties
- items waarbij maar een scorecategorie geobserveerd is (alles goed of alles fout).

Als deze opgaven zijn verwijderd voeren we een initiële 2PL-kalibratie uit waarna we items uit de dataset verwijderen met een a -parameter van minder dan 0.1.

Alle ankeritems worden gecontroleerd op mogelijk differentieel functioneren (DIF) tussen aanbieders onderling, maar ook in vergelijking met afnames uit het verleden. Na het uitschakelen of loskoppelen van items voeren we bovenstaande procedure opnieuw uit. De bevindingen rapporteren we terug aan de aanbieders. In deze bevindingen staan:

- de itemparameters van de nieuwe ankeritems en/of eventueel opnieuw gekalibreerde ankeritems
- eventuele losgekoppelde dan wel uitgeschakelde ankeritems
- voor de reguliere items of ze de grens van $a < 0.1$ niet halen (dit kan immers wijzen op sleutelfouten of andere problemen).

Uitgeschakelde of losgekoppelde items zullen voorzien worden van een toelichting en plaatjes met karakteristieke itemcurves. De kalibratie wordt uitgevoerd met behulp van het R-script `kalibreer_anker.R`.

4.1 TERUGLEVERING ITEMPARAMETERS GEZAMENLIJK ANKER

Stichting Cito levert de aanbieders een csv-bestand met gekalibreerde itemparameters, zowel van de bestaande als de nieuwe items 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_2024.csv`

Tevens levert Stichting Cito dan aan de aanbieders een bestand met uitgeschakelde items die een te lage 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`

5. NORMERINGSPROCEDURE IRT

In dit hoofdstuk wordt aangegeven welke stappen genomen worden om tot de normering te komen.

² <https://github.com/dexter-psychometrics/dexterMML>



Dit betreft de normering van de referentieniveaus en de toetsadviezen.

5.1 IRT-NORMERING

De IRT-normering wordt uitgevoerd door middel van een fixatie op de itemparameters van het gezamenlijk anker. Dit gebeurt voor iedere toetsaanbieder apart, waarbij door Stichting Cito net als bij de kalibratie van het gezamenlijk anker (zie hoofdstuk 4 *Kalibratie gezamenlijk ankeritems*) alleen items meegenomen worden die aan drie eisen voldoen: (1) meer dan 200 observaties; (2) meer dan 1 gebruikte scorecategorie; en (3) in het 2PL een a-parameter van 0.1 of hoger. In de kalibratie worden dezelfde MML-populaties gebruikt als bij de kalibratie van het gezamenlijk anker, en alleen leerlingen uit regulier BO worden meegenomen. Zowel de itemparameters als de populatieparameters worden in de kalibratie geschat. Alle stappen van de IRT-normering worden uitgevoerd met behulp van het R-script `irt_normering.R`.

Deze opgeleverde *toetseigen* itemparameters en populatieparameters, evenals de daaruit volgende percentages behaalde referentieniveaus en toetsadviezen dienen slechts ter controle. Om de eigen toets te normeren kunnen de toetsaanbieders niet zonder meer de itemparameters zoals opgeleverd door Stichting Cito gebruiken. Bij een CAT heeft Stichting Cito bijvoorbeeld niet de beschikking over alle itembankgegevens om precieze schattingen te maken van de itemparameters van items met weinig observaties. Een toetsaanbieder zal dus zelf na moeten gaan of het over voldoende gegevens beschikt om betrouwbaar te normeren, en in het geval van een itembank na de afname een nieuwe bankkalibratie moeten uitvoeren met de juiste fixatie op de schaal van het gezamenlijk anker. De globale cesuurpunten voor de geharmoniseerde IRT-normering zoals beschreven in dit normerings-handboek kunnen wel gewoon toegepast worden.

5.2 REFERENTIENIVEAUS

Nadat de itemparameters geschat zijn, worden voor de drie verplichte onderdelen vaardigheidsschattingen uitgevoerd voor iedere leerling met behulp van Warm's gemiddelde gewogen waarschijnlijkheidsschatter. Deze vaardigheidsschattingen worden vervolgens vergeleken met de cesuurpunten voor ieder van de referentieniveaus:

Tabel 1: cesuurpunten voor de referentieniveaus in de IRT-normering.

referentieniveau	cesuur 2PL
Lezen 1F	-2.12085
Lezen 2F	-0.81552
Rekenen 1F	-1.59269
Rekenen 1S	0.02194
Taalverzorging 1F	-2.00258
Taalverzorging 2F	-0.43391

Van iedere aanbieder zal de proportie referentieniveau behaald uitgerekend worden.

5.3 TOETSADVIEZEN³

De cesuurpunten voor de toetsadviezen zijn gedefinieerd op een schaal met gemiddelde latente vaardigheden (GLV), genomen over de drie verplichte onderdelen. De wegingsfactoren van de verschillende onderdelen in dit gemiddelde zijn:

Tabel 2: wegingsfactoren van de onderdelen voor de GLV.

lezen	rekenen	taal
0.3000	0.4500	0.2500

De GLV's worden vervolgens vergeleken met cesuurpunten om iedere leerling op basis hiervan een toetsadvies toe te kennen:

³ Naast de verplichte onderdelen kunnen ook eventuele additionele en optionele onderdelen die meetellen in het toetsadvies van iedere aanbieder gekalibreerd worden. Geen van de toetsaanbieders biedt voor 2024 optionele onderdelen aan die meetellen voor het toetsadvies.

Tabel 3: cesuurpunten voor de toetsadviezen op de GLV-schaal in de IRT-normering.

grenspunt	cesuur 2PL
pro/vmbo bb -> vmbo bb/kb	-2.16753
vmbo bb/kb -> vmbo kb/gl-tl	-1.13056
vmbo kb/gl-tl -> vmbo gl-tl/havo	-0.48615
vmbo gl-tl/havo -> havo/vwo	0.18100
havo/vwo -> vwo	0.79278

5.4 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 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

5.5 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

5.6 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

5.7 TERUGLEVERING VAARDIGHEIDSSCHATTINGEN

Bevat voor iedere leerling de vaardigheid op ieder onderdeel in het 2PL-model, geschat via Warm's gemiddelde gewogen waarschijnlijkheid. 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	Schooltype, zelfde indeling als leerlinggegevens
theta_[onderdeel]	Vaardigheidsschatting voor ieder onderdeel
glv	GLV
toetsadvies	Toetsadvies volgens de IRT-normering

Als bestandsnaam wordt gebruikt:

TOETS_vaardigheden.csv

5.8 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

5.9 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

5.10 TOEPASSING NORMERING TEN BEHOEVE VAN LEERLINGRAPPORTAGES

De teruggeleverde resultaten vormen slechts een normeringsadvies op basis van de aangeleverde data van de huidige afname met fixatie op de schaal van het gezamenlijk anker. Om de IRT-normering toe te passen voor de leerlingrapportages kunnen niet zonder meer de itemparameters zoals teruggeleverd gebruikt worden. Een toetsaanbieder beschikt mogelijk over meer waarnemingen van de items dan geleverd zijn aan het CvTE. In dat geval zal een nieuwe kalibratie door de toetsaanbieder de nauwkeurigheid vergroten. Daarnaast worden bij een CAT bijvoorbeeld niet alle itembankgegevens gedeeld, die noodzakelijk zijn om precieze schattingen te maken van de itemparameters van items met weinig observaties. Een toetsaanbieder zal dus zelf na moeten gaan of het over voldoende gegevens beschikt om betrouwbaar te normeren, en in het geval van een itembank na de afname een nieuwe bankkalibratie moeten uitvoeren met de juiste fixatie op de schaal van het gezamenlijk anker. Bij een eigen kalibratie met parameterfixatie op de schaal van het gezamenlijk anker kunnen dan vervolgens de in dit handboek genoemde cesuurpunten toegepast worden.

6. NORMERING BIJ INZET DOORSTROOMTOETS OCW

Omdat een mogelijk inzetscenario van de doorstroomtoets OCW het uitlekken van gezamenlijk ankeritems is, bevat de doorstroomtoets OCW geen gezamenlijk ankeritems. In plaats van een parameterfixatie op het gezamenlijk anker, worden de vaardigheidsschalen in de normering van de doorstroomtoets OCW gefixeerd door middel van een koppeling naar de Centrale Eindtoets. De



doorstroomtoets OCW bestaat namelijk uit items 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 doorstroomtoets OCW 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 dit handboek toegepast worden om tot de waarden voor leerlingrapportages te komen.

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

1. Een kleinschalige inzet vanwege een gelokaliseerde calamiteit, minder dan 1000 leerlingen.
2. Een grootschalige inzet met hoge leerlingaantallen.

Bij een kleinschalige inzet zijn er te weinig observaties van de items uit de doorstroomtoets OCW voor een betrouwbare kalibratie. In dat geval zal de doorstroomtoets OCW 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 calamiteiten-team.

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

7 DATALEVERING TEN BEHOEVE VAN HET TOELATINGS- EN DOORSTROOMONDERZOEK

7.1 ACHTERGROND

Voor de normering moet o.a. toelatings- en doorstroomonderzoek (hierna: doorstroomonderzoek) verricht worden. Doorstroomonderzoek 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 toets 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 voortgezet onderwijs.

7.2 PROCES VERWERKING PERSOONSgegevens

Het doorstroomonderzoek voert Stichting Cito uit voor het CvTE. Hierbij maakt Stichting Cito gebruik van een gegevensbestand dat zijzelf samenstellen met gegevens van DUO uit het Register Onderwijs-deelnemers (ROD) en gegevens die afkomstig zijn van de toetsaanbieders. De gegevens die Stichting Cito nodig heeft, komen dus uit twee verschillende bronnen. Voor het onderzoek is het noodzakelijk dat de bestanden met gegevens van leerlingen van een aanbieder worden gekoppeld aan de gegevens van die leerling 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 het gebruik van persoonsgegevens tot een minimum te beperken zal gebruik gemaakt worden van een pseudoniem per individuele leerling in de uitwisseling tussen enerzijds toetsaanbieders en DUO en anderzijds Stichting Cito. Het doel van dit pseudoniem is dus *enkel* het kunnen koppelen van de gegevens van individuele leerlingen afkomstig van de toetsaanbieders en DUO, door Stichting Cito.

Door middel van het toepassen van een reeds bestaand pseudonimiseringsalgoritme kunnen toetsaanbieders en DUO *zelf* een 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. Het pseudoniem per individuele leerling komt voor DUO en de toetsaanbieders in de databestanden dan overeen, en deze databestanden met pseudoniemen worden vervolgens gedeeld met Stichting Cito, waardoor Stichting Cito ten behoeve van het onderzoek de bestanden samen kan voegen. Er is 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 aanbieders die de gegevens hebben geleverd kunnen het pseudoniem weer koppelen aan de persoonsgegevens waarmee het pseudoniem is gegenereerd. Op die manier beschikt Stichting Cito onder verantwoordelijkheid van het



CvTE over zo min mogelijk gevoelige gegevens om deze wettelijke taak goed uit te kunnen voeren.

7.3 SPECIFICATIE BESTANDLEVERING

Datalevering als CSV

- Scheidingstekens ';' ,
- Teksten zonder aanhalingstekens
- Decimaal scheidingstekens: komma
- 1^e regel bevat kolomnamen

Levering door DUO:

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

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

Pseudoniem	Zie specificatie CvTE
Toetssoort	Uit waardelijst 'Toetssoort'
Ontheffing	J/N
Toetsscore	
Kalenderjaar	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'

Levering door aanbieders

Naamgeving [Toetssoort]-dso-' + [YYYYMMDD] + '.csv'

- [Toetssoort] = uit waardelijst 'Toetssoort'
- [YYYYMMDD] = datum aanmaak bestand

Pseudoniem	Zie hoofdstuk 3.2.3
Toetssoort	Uit waardelijst 'Toetssoort'
Afnamemodus	Uit waardelijst 'Afnamemodus'
Toetsscore	
Kalenderjaar	YYYY
Referentieniveau rekenen	Uit waardelijst 'Referentieniveau Rekenen'
Referentieniveau lezen	Uit waardelijst 'Referentieniveau Taal'
Referentieniveau taalverzorging	Uit waardelijst 'Referentieniveau Taal'
Score GLV	de GLV ¹ is zoals bepaald door de aanbieder (dus inclusief eventuele optionele onderdelen die meetellen voor het toetsadvies)
Score lat.v.s. rekenen	Score lat.v.s. Rekenen ¹
Score lat.v.s. lezen	Score lat.v.s. Lezen ¹
Score Lat.v.s. taalverzorging	Score Lat.v.s. Taalverzorging ¹
Toetsadvies	Uit waardelijst 'Toetsadvies'

¹ De GLV en de lat. vaardigheid bij voorkeur in het 2PL

Waardelijsten

Toetssoort	Afnamemodus	Referentieniveaus rekenen	Referentieniveau taal
ROUTE_8	papier	L1F	L1F
ICE	digitaal	1F	1F
DIA		1S	2F
AMN			
LEERLING_IN_BEELD			
DOE			
DOORSTROOMTOETS OCW			



Toetssoort	Afnamemodus	Referentieniveaus rekenen	Referentieniveau taal
------------	-------------	------------------------------	-----------------------

0011

NiveauVO	Toetsadvies	Schooladvies
pro	pro/vmbo bb	pro
vmbo bb	vmbo bb/vmbo kb	vmbo bb
vmbo kb	vmbo kb/vmbo gl-tl	vmbo bb/vmbo kb
vmbo gl	vmbo gl-tl/havo	vmbo kb
vmbo gl-tl	havo/vwo	vmbo kb/vmbo gt-tl
vmbo tl	vwo	vmbo gl-tl
havo		vmbo gt-tl/havo
vwo		havo
		havo/vwo
		vwo

Type onderwijs
BO
SBO
SO

Gegevensverwerkingen

Gegevensverwerking	Persoonsgegevens
1. DUO verzamelt gegevens uit het ROD.	1. Achternaam leerling 2. Geboortedatum leerling 3. Instellingscode (BRIN4) 4. Toetsscore leerling 5. Kalenderjaar doorstroomtoets 6. Toets ID 7. Referentieniveaus 8. Voorlopige en definitieve schooladvies VO door school 9. Verblijfsjaar 10. Leerjaar 11. Niveau 12. Ontheffing 13. Type onderwijs
2. Toetsaanbieders verzamelen en genereren gegevens.	1. Achternaam leerling 2. Geboortedatum leerling 3. Instellingscode (BRIN4) 4. Toetsscore leerling 5. Kalenderjaar doorstroomtoets 6. Toetssoort 7. Referentieniveaus 8. Score op de gewogen latentvaardigheidsschaal 9. Scores op de latentvaardigheidsschaal van de losse terreinen 10. Toetsadvies 11. Afnamemodus
3. DUO en toetsaanbieders genereren een pseudoniem met een pseudonimiserings-algoritme.	1. Instellingscode (BRIN4) 2. Achternaam leerling 3. Geboortedatum leerling 4. Toetsscore leerling
4. DUO levert gegevens aan Stichting Cito.	1. Pseudoniem 2. Toets ID 3. Toetsscore leerling 4. Kalenderjaar doorstroomtoets 5. Referentieniveaus 6. Verblijfsjaar 7. Leerjaar 8. Niveau 9. Voorlopige en definitieve schooladvies VO door school 10. Ontheffing 11. Type onderwijs



Gegevensverwerking	Persoonsgegevens
5. Toetsaanbieders leveren gegevens aan Stichting Cito.	1. Pseudoniem 2. Toetssoort ¹ 3. Toetsscore leerling ² 4. Kalenderjaar doorstroomtoets ³ 5. Referentieniveaus ⁴ 6. Score op de gewogen latentevaardigheidsschaal ⁵ 7. Scores op de latentevaardigheidsschaal van de losse terreinen ⁶ 8. Toetsadvies 9. Afnamemodus ⁷
6. Stichting Cito koppelt de gegevens en voert onderzoek uit.	1. Pseudoniem 2. Toets ID 3. Toetsscore leerling 4. Kalenderjaar doorstroomtoets 5. Referentieniveaus 6. Score op de gewogen latentevaardigheidsschaal 7. Scores op de latentevaardigheidsschaal van de losse terreinen 8. Schooladviezen VO 9. Verblijfjaar 10. Leerjaar 11. Niveau 12. Toetsadvies 13. Ontheffing 14. Afnamemodus 15. Voorlopig en definitief schooladvies 16. Type onderwijs

- ¹ Waardenlijst Toetssoort:
ROUTE_8 – Route 8
ICE – IEP Doorstroomtoets
DIA – Dia-doorstroomtoets
AMN – AMN Doorstroomtoets
LEERLING_IN_BEELD – Leerling in beeld-doorstroomtoets
DOE – Doe overheidsdoorstroomtoets
Doorstroomtoets OCW – Doorstroomtoets OCW
CET – Centrale eindtoets
- ² Zie tabel op pagina 3.
- ³ Het kalenderjaar waarin de doorstroomtoets is afgenomen.
- ⁴ L1F, 1F, 2F of 1S
- ⁵ Op 3 decimalen afgerond.
- ⁶ Op 3 decimalen afgerond.
- ⁷ Papier of digitaal

Toelichting:

Hoewel het op het eerste gezicht lijkt alsof DUO en toetsaanbieders beide dezelfde persoonsgegevens leveren, zit er toch een verschil tussen beide gegevenssets. DUO beschikt in het ROD namelijk niet over de scores die leerlingen op elk toetsonderdeel behalen, terwijl deze gegevens wel nodig zijn voor het uitvoeren van het doorstroomonderzoek. De toetsaanbieders beschikken hier wel over. Daarentegen beschikken de toetsaanbieders niet over de doorstroomgegevens (Verblijfjaar, Leerjaar, Niveau) maar DUO wel.