



> Retouradres Postbus 315 3500 AH Utrecht

## Stichting Cito

### Bureau van het CvTE

Muntstraat 7  
3512 ET Utrecht  
Postbus 315  
3500 AH Utrecht  
Nederland  
[www.cvte.nl](http://www.cvte.nl)

### Contactgegevens

030 - 28 40 700  
[info@cvte.nl](mailto:info@cvte.nl)

Datum 4 september 2023  
Betreft Besluit aanvraag erkenning DOE overheidsdoorstroomtoets

---

### Onze referentie

CvTE-23.00939

### Bijlagen

(1) Advies Stichting Cito  
(2) Advies Adviescommissie  
Toetsen PO

Geachte,

Met deze brief besluit het College voor Toetsen en Examen (CvTE) naar aanleiding van uw verzoek tot erkenning van de DOE overheidsdoorstroomtoets. Uw verzoek is ontvangen op 31-05-2023 (door ons geregistreerd met correspondentienummer: 38129).

## 1 BESLUIT EN WETTELIJK KADER

Gelet op artikel 3a, eerste lid onder a, van de Wet College voor toetsen en examens, artikel 45b, derde lid, Wet op het primair onderwijs, de artikelen 4 en 9, tweede en derde lid, Toetsbesluit PO, heeft het College besloten de DOE overheidsdoorstroomtoets te erkennen voor vier jaar met ingang van het schooljaar 2023-2024.

## 2 MOTIVERING

Bij de beoordeling van de DOE overheidsdoorstroomtoets heeft het College beoordeeld of de toets voldoet aan de wettelijke vereisten en de Regeling beoordelingskader doorstroomtoets po. Het College heeft hierbij gebruik gemaakt van het advies van Stichting Cito. De Stichting heeft bij haar advisering de Regeling beoordelingskader doorstroomtoets po toegepast. Stichting Cito heeft op basis van de onderwijskundige, organisatorische en psychometrische aspecten geadviseerd de toets te erkennen (zie bijlage 1). Het College heeft zich er van vergewist dat het advies van de Stichting op zorgvuldige wijze tot stand is gekomen (zie bijlage 2). Het College neemt het advies van Stichting Cito over.

### 3 TOT SLOT

Dit besluit en bijbehorende adviezen zullen zowel in de Staatscourant als op de website van het CvTE (<https://www.cvtetoetsenpo.nl>) worden gepubliceerd.

Hoogachtend,

drs. J.H. van der Vegt MPM  
voorzitter

*Tegen dit besluit kunt u binnen zes weken na de dag waarop dit is bekend gemaakt een bezwaarschrift indienen. Het bezwaarschrift moet door de indiener zijn ondertekend en bevat ten minste zijn naam en adres, de dagtekening, een omschrijving van het besluit waartegen het bezwaar is gericht en de gronden waarop het bezwaar rust. Dit bezwaarschrift moet worden gericht aan: de voorzitter van het College voor toetsen en examens, Postbus 315, 3500 AH Utrecht.*

## Advies doorstroomtoets

### DOE overheidsdoorstroomtoets

Stichting Cito – 29-8-2023

Ter beoordeling van de DOE overheidsdoorstroomtoets geeft Stichting Cito advies aan het College voor Toetsen en Examens (CvTE). Het advies is tot stand gekomen aan de hand van de [regeling beoordelingskader doorstroomtoets PO](#).

Dit advies betreft een: vierjaarlijkse erkenning

Type doorstroomtoets: lineair (digitaal) en MST

Het CvTE is met Stichting Cito, aanbieder van de DOE overheidsdoorstroomtoets, de volgende coulance-afspraken (referentienummer 23.00443) overeengekomen: Stichting Cito kan opgaven op 31 mei 2023 nog niet aanbieden in de definitieve lay-out in de afnamespeler en streeft naar de volgende wijze van aanbieden ter beoordeling:

- Weergave in een klikbare interactie van alle opgaven in een default lay-out van de afnamespeler. Hierdoor kunnen beoordelaars een goed beeld krijgen van de interactie en de opgaven.
- Print-screens van specifieke opgavenvormen in definitieve lay-out. Hierdoor kunnen beoordelaars een goed beeld krijgen van de scherm lay-out die wordt ontwikkeld.

Is het beoordelingsteam gedurende het beoordelingsproces tegen problemen aangelopen?

Ja

Toelichting indien ja: bij de initieel aangeleverde informatie miste het beoordelingsteam een adequate verantwoording voor de beoordeling van kwaliteitseis KA6. Tijdens het hoor- en wederhoor gesprek is dit besproken. De aanbieder heeft op basis van het gesprek aanvullende informatie aangeleverd, die vervolgens is beoordeeld.

## 1. Aanvraag en proces

Op 2 juni 2023 heeft Stichting Cito de documenten voor de beoordeling van de DOE overheidsdoorstroomtoets ontvangen. De door de aanbieder ingevulde leeswijzer maakt inzichtelijk op welke plek in de documentatie de kwaliteitscriteria worden verantwoord. De leeswijzer maakt een snelle controle op volledigheid mogelijk. Op 2 juni 2023 zijn de documenten gedeeld met het beoordelingsteam. De DOE overheidsdoorstroomtoets is beoordeeld door een team bestaande uit één expert op het gebied van taal, één expert op het gebied van rekenen en één expert op het gebied van psychometrie. Daarnaast waren twee editors van Stichting Cito betrokken bij de beoordeling, waarvan één onderwijskundig editor en één psychometrisch editor. Op 30 juni 2023 hebben de beoordelaars en editors overleg gehad over de bevindingen. De editors hebben de uitkomst van dit overleg verwerkt tot een conceptadvies. Dit conceptadvies is met de aanbieder gedeeld in aanloop naar het hoor- en wederhoor gesprek op 11 juli 2023.

## 2. Algemene informatie

De algemene informatie over de DOE overheidsdoorstroomtoets komt uit de wetenschappelijke verantwoording.

De DOE overheidsdoorstroomtoets heeft als doel om (1) een onafhankelijk toetsadvies te geven en (2) om te rapporteren op de beheersing van de referentieniveaus. De overheidsdoorstroomtoets bestaat uit de toetsonderdelen Nederlandse taal (Lezen en Taalverzorging) en Rekenen. De toetsscore van een leerling wordt gebaseerd op een weging van de vaardigheidsscores van de leerling op de drie verschillende onderdelen Lezen (30%), Taalverzorging (25%) en Rekenen (45%). De toetsscore wordt vervolgens omgezet naar een toetsadvies. De toetsaanbieder gebruikt de voorgeschreven indeling van de zes toetsadviescategorieën. Uit de overheidsdoorstroomtoets volgt ook een rapportage op de referentieniveaus Taalverzorging (1F en 2F), Rekenen (1F en 1S) en Lezen (1F en 2F).

### ***Inhoudelijke verantwoording***

De DOE overheidsdoorstroomtoets is in eerste instantie een adaptieve, multi-stage toets (MST). Dit wil zeggen dat de toets bestaat uit verschillende toetspaden. Een leerling doorloopt de toets op basis van de behaalde score op elk toetsblok. Elk toetsblok bestaat uit een vaste hoeveelheid opgaven (d.w.z., voor elke leerling dezelfde hoeveelheid opgaven) en leerlingen kunnen binnen een toetsblok terugbladeren. In het design is de dekking van de inhoudsdoelen en de spreiding van de referentieniveaus geborgd. Naast de MST versie wordt de DOE ook als lineaire digitale toets aangeboden. De lineaire versie van de toets is samengesteld met de toetsblokken van het gemiddelde toetspad van de MST versie.

De aanbieder heeft bij indiening de volgende documentatie aangeleverd:

- Indieningsbrief
- Leeswijzer
- Wetenschappelijk verantwoording
- Document instructie voor beoordelaars
- Overzicht samenstelling verantwoordingstoetsen
- Stappenplan inschrijving
- Leeswijzer rapportages
- Toets reglement

- Calamiteitenplan
- ISO27001 certificaat + verklaring van toepassing

### 3. Proces beoordeling kwaliteitscriteria

Hieronder volgt een kort overzicht van de belangrijkste bevindingen van het beoordelingsteam op de drie verschillende aspecten. Een overzicht van de beoordeling per kwaliteitscriterium is weergegeven in bijlage 1.

#### Onderwijskundige aspecten

De aanbieder toont in de verantwoording op kwaliteitscriteria 3.2 tot en met 3.6 aan dat de toets voldoet aan de gestelde eisen. De opgaven zijn voorzien van passende itemkenmerken en item-labels. Het beoordelingsteam heeft vastgesteld dat de domeindekking in de verschillende routes binnen de MST op orde is. Ook zal een leerling, onafhankelijk van de route die de leerling doorloopt, voldoende opgaven met verschillende referentieniveaus tegenkomen. Het beoordelingsteam stelt dat het rapporteren op behaald referentieniveau hierdoor voor elke leerling te onderbouwen is. De lineaire toets is evenwichtig opgebouwd en laat een goede verdeling van referentieniveaus zien. Het beoordelingsteam heeft de opgaven in een tijdelijke afname-omgeving bekeken. In de wetenschappelijke verantwoording zijn screenshots van opgaven in de toekomstige lay-out opgenomen. Dit is conform de gemaakte coullance-afspraken. Het beoordelingsteam heeft door de screenshots een goed beeld gekregen van hoe de opgaven uiteindelijk in de toets worden getoond.

#### Organisatorische aspecten

De kwaliteitseisen met betrekking tot de organisatorische aspecten zijn duidelijk verantwoord. De DOE overheidsdoorstroomtoets biedt diverse mogelijkheden voor leerlingen met specifieke ondersteuningsbehoeften. De verantwoordingsdocumenten rondom de rapportage zijn compleet en begrijpelijk. De aanbieder is ISO 27001 gecertificeerd. De kwaliteitscriteria RB1 tot en met RB4.2 zijn daarom niet van toepassing.

#### Psychometrische aspecten

Omdat de DOE overheidsdoorstroomtoets een nieuwe toets is, is de rapportage op afname (5.3) niet van toepassing. De steekproefgrootte voldoet aan de criteria voor nieuwe aanbieders en is gebruikt om de kwaliteit van opgaven te bepalen ten behoeve van de selectie. De toetspopulatie en pretestpopulatie konden niet worden vergeleken, omdat het gaat om een nieuwe toets. Daarom is de pretestpopulatie vergeleken met de landelijke populatie, en in geval van discrepanties met de populatie van de Centrale Eindtoets. Deze populaties komen in grote mate overeen. In de verantwoording op kwaliteitseis KA6 miste het beoordelingsteam bij de initiële indiening een toelichting over de unidimensionaliteit van de data. De aanbieder heeft verschillende DIF analyses uitgevoerd om de passing van het model te tonen, maar er ontbrak een expliciete verantwoording op het gebied van unidimensionaliteit. Dit kwam ter sprake tijdens het hoor- en wederhoorgesprek.

### 4. Hoor- en wederhoorgesprek

Op 11 juli 2023 heeft onder regie van het CvTE het hoor- en wederhoorgesprek plaatsgevonden. Vanuit het beoordelingsteam kwam kwaliteitseis KA6 ter sprake, omdat het team informatie miste om deze eis te kunnen beoordelen. In het gesprek heeft de aanbieder een toelichting gegeven hoe de unidimensionaliteit van de toetsonderdelen is onderzocht. Er is afgesproken dat deze toelichting schriftelijk met het CvTE zou worden gedeeld.

Het beoordelingsteam was ook benieuwd in hoeverre een andere afname-omgeving invloed heeft op de geschatte itemparameters. De aanbieder heeft toegelicht dat als de afnameomgeving gevolgen heeft voor de psychometrische gegevens, dit niet problematisch is, omdat er achteraf wordt genormeerd.

#### 5. Proces richting definitief advies

De op 11 juli 2023 mondeling gegeven toelichting is op 14 juli 2023 schriftelijk aangeleverd. De informatie is door de editors van Stichting Cito bekeken. De aanbieder toont aan dat er latente correlaties tussen de toetsonderdelen zijn berekend. Hieruit blijkt dat de toetsonderdelen voldoende unidimensioneel zijn. Hiermee is de kwaliteitseis KA6 voldoende verantwoord.

#### 6. Oordeel en advies

Met betrekking tot de in het beoordelingskader gestelde kwaliteitseisen concludeert Stichting Cito als volgt:

Aspect	Paragraaf	Kwaliteitseis	Oordeel	
Onderwijskundig	3.2	Inhoudsvaliditeit doorstroomtoets	Voldoende	
	3.3	Verantwoording	Voldoende	
	3.4.1	Wettelijk verplicht domein: Lezen	Voldoende	
	3.4.2	Wettelijk verplicht domein: Taalverzorging	Voldoende	
	3.4.3	Optioneel domein: Schrijven	n.v.t.	
	3.4.4	Optioneel domein: mondelinge taalvaardigheid	n.v.t.	
	3.4.5	Optioneel domein: Begrippenlijst	n.v.t.	
	3.5.1	Verdeling toetsopgaven over domeinen in toetsmatrijs en toetssamenstelling	Voldoende	
	3.5.2	Verdeling toetsopgaven over onderdelen	Voldoende	
	3.5.3	Opgaven met en zonder context	Voldoende	
	3.5.4	Gebruik kladpapier	Voldoende	
	3.5.5	Gebruik rekenmachine	Voldoende	
	3.6	Toetsopgaven Nederlandse taal en rekenen op niveau 1F 2F/1S	Voldoende	
	Organisatorisch	4.1	Leerling rapport	Voldoende
		4.2	Leerlingen	Voldoende
4.3		Afname doorstroomtoets	Voldoende	
4.4		Beveiligingsaspecten	Voldoende	
Psychometrisch	5.2.1	Steekproefkader en samenstelling steekproef	Voldoende	
	5.2.2	Kalibratie en kwaliteit van items	Voldoende	
	5.2.3	Toetssamenstelling	Voldoende	
	5.3	Rapportage achteraf	n.v.t.	

Stichting Cito geeft op basis van de verstrekte informatie ter verantwoording van de DOE overheidsdoorstroomtoets een positief advies voor de vierjaarlijkse erkenning van de DOE overheidsdoorstroomtoets.

Dit advies is vastgesteld op 29 augustus 2023 te Arnhem.

## Bijlage 1: Toelichting beoordeling per kwaliteitscriterium

### 3 Onderwijskundige aspecten

#### 3.2 Inhoudsvaliditeit doorstroomtoets

Code	Vragen	Toelichting	Antwoord
V1	Is de toetsmatrijs een adequate representatie van het meetdoel voor in ieder geval de wettelijk verplichte domeinen van Nederlandse taal en rekenen én de, indien van toepassing, optionele domeinen van Nederlandse taal?	De toetsmatrijs is een adequate representatie van de meetdoelen zoals beschreven in het referentiekader. De verdeling van de domeinen is evenwichtig. In tabel 1 van de wetenschappelijke verantwoording wordt de verdeling van opgaven over de domeinen inzichtelijk gemaakt. Het document 'overzicht samenstelling verantwoordingstoetsen' laat de spreiding van onderdelen per toetsdomein zien en toont een adequate representatie van het meetdoel in beide toetsversies.	Ja
V2	Voldoen de toetsopgaven van enerzijds de wettelijk verplichte domeinen van Nederlandse taal en rekenen én anderzijds de, indien van toepassing, optionele domeinen van Nederlandse taal in de toets aan de kwaliteitscriteria voor de inhoudsvaliditeit (zijnde de relevantie van de inhoud van de toets) van toetsopgaven?	De toetsopgaven voldoen aan de criteria van de inhoudsvaliditeit. Op pagina 13 en 14 van de wetenschappelijke verantwoording wordt het constructieproces beschreven. De opgaven voldoen aan de vijf observatie-categorieën.	Ja
V3	Zijn de toetsopgaven van wettelijk verplichte domeinen van Nederlandse taal en rekenen én de, indien van toepassing, optionele domeinen van Nederlandse taal in de toets correct geconstrueerd?	De opgaven zijn correct geconstrueerd. De opgaven voldoen aan de in de checklist beschreven voorwaarden. Het beoordelingsteam heeft, conform de coulance-afpraak, de opgaven in een tijdelijke afname-omgeving bekeken. De vormgeving en de verwachte functionaliteit van de nieuwe omgeving ziet er op basis van de aangeleverde screenshots volgens het beoordelingsteam goed uit.	Ja

#### 3.3 Verantwoording

Code	Vraag	Toelichting	Antwoord
H1	Is er (een set van) verantwoordingsdocumenten aangeboden aan het CvTE met daarin voor wat betreft in ieder geval de wettelijk verplichte domeinen van Nederlandse taal en rekenen én, indien van toepassing, de optionele domeinen van Nederlandse taal een verantwoording van de gemaakte keuzes ten aanzien van de toetsinhoud, de afnamevorm, de psychometrische aspecten en de kwaliteitszorg rondom de itemconstructie van de toets?	De gevraagde verantwoordingsdocumenten zijn aangeleverd. De wetenschappelijke verantwoording beschrijft de toetsinhoud (o.a. pagina 11 tot en met 13) en de psychometrische onderbouwing (Hoofdstuk 5) van de toets. Het toetsreglement is volledig en duidelijk. In de verantwoording wordt aangegeven dat alle onderdelen meetellen bij het berekenen van het toetsadvies. De zaai-opgaven worden niet meegenomen bij deze berekening.	Ja

#### 3.4 Terrein Nederlandse taal

##### 3.4.1 Lezen

###### Subdomein Zakelijke Teksten

Code	Vragen	Toelichting	Antwoord
L1.1	Lezen de leerlingen informatieve teksten en maken ze daar opdrachten bij?	In tabel 6 van de wetenschappelijke verantwoording (pagina 19) wordt een overzicht gegeven van het aantal opgaven per toetspad over de verschillende soorten zakelijke teksten. Leerlingen lezen informatieve teksten en maken hier opgaven bij.	Ja
L1.2	Lezen de leerlingen instructieve teksten en maken ze daar opdrachten bij?	In tabel 6 van de wetenschappelijke verantwoording (pagina 19) wordt een overzicht gegeven van het aantal opgaven per toetspad over de verschillende soorten zakelijke teksten. Leerlingen lezen instructieve teksten en maken hier opgaven bij.	Ja
L1.3	Lezen de leerlingen betogende teksten en maken ze daar opdrachten bij?	In tabel 6 van de wetenschappelijke verantwoording (pagina 19) wordt een overzicht gegeven van het aantal opgaven per toetspad over de verschillende soorten zakelijke teksten. Leerlingen lezen betogende teksten en maken hier opgaven bij.	Ja
L2	Voldoen de teksten die de leerlingen lezen aan de tekstenmerken inhoud (thema), vorm (taalgebruik en structuur) en bedoeling (doel), zoals gespecificeerd in Checklist 1 van deze regeling en zoals nader omschreven in het Referentiekader?	In de verantwoording staan op pag. 7-8 een aantal kenmerken waar teksten aan voldoen. De aanbieder heeft gebruikgemaakt van UXkids, een organisatie die is gericht op het gebruiksvriendelijk ontwikkelen van software voor kinderen. De adviezen van UXkids zijn terug te zien in het onderdeel Lezen. Vooral de 1F teksten hebben een lage informatiedichtheid en bevatten weinig laag frequente woorden.	Ja

Code	Vragen	Toelichting	Antwoord
L3.1	Wordt in de toets als kenmerk van de taakuitvoering techniek en woordenschat getoetst?	In tabel 7 van de wetenschappelijke verantwoording (pagina 21) wordt een overzicht gegeven van de verdeling van de verschillende taakuitvoeringen. Techniek en woordenschat wordt getoetst.	Ja
L3.2	Wordt in de toets als kenmerk van de taakuitvoering begrijpen getoetst?	In tabel 7 van de wetenschappelijke verantwoording (pagina 21) wordt een overzicht gegeven van de verdeling van de verschillende taakuitvoeringen. Begrijpen wordt getoetst.	Ja
L3.3	Wordt in de toets als kenmerk van de taakuitvoering interpreteren getoetst?	In tabel 7 van de wetenschappelijke verantwoording (pagina 21) wordt een overzicht gegeven van de verdeling van de verschillende taakuitvoeringen. Interpreteren wordt getoetst.	Ja
L3.4	Wordt in de toets als kenmerk van de taakuitvoering samenvatten getoetst?	In tabel 7 van de wetenschappelijke verantwoording (pagina 21) wordt een overzicht gegeven van de verdeling van de verschillende taakuitvoeringen. Samenvatten wordt getoetst.	Ja
L3.5	Wordt in de toets als kenmerk van de taakuitvoering opzoeken getoetst?	In tabel 7 van de wetenschappelijke verantwoording (pagina 21) wordt een overzicht gegeven van de verdeling van de verschillende taakuitvoeringen. Opzoeken wordt getoetst.	Ja
L3.6	Wordt in de toets als kenmerk van de taakuitvoering evalueren getoetst (conform de Toetswijzer PO is het onderwerp 'evalueren' een optioneel en niet verplicht onderdeel van de toets)?	De taakuitvoering evalueren wordt niet getoetst in de DOE overheidsdoorstroomtoets. Evalueren is een optioneel onderdeel. Hierdoor is kwaliteitseis L3.6 voor de DOE overheidsdoorstroomtoets niet van toepassing.	n.v.t.
L3.7	Worden in de toets de verschillende kenmerken van de taakuitvoering evenwichtig getoetst binnen het subdomein zakelijke teksten?	In tabel 7 (pag. 21 van de wetenschappelijke verantwoording) staat een overzicht van kenmerken van de taakuitvoering. In het Excel-document staat elke vraag, inclusief kenmerk/label. De taakuitvoeringen worden in beide toetsversies evenwichtig getoetst.	Ja

#### Subdomein Fictionele, narratieve en literaire teksten

Code	Vragen	Toelichting	Antwoord
L4	Lezen de leerlingen fictionele, narratieve en/of literaire teksten en maken ze daar opdrachten bij?	In tabel 7 van de wetenschappelijke verantwoording (pagina 21) wordt een overzicht gegeven van de verdeling van de verschillende taakuitvoeringen. De DOE Overheidsdoorstroomtoets bevat verhalende teksten en teksten met poëzie. Leerlingen maken opgaven bij de teksten en gedichten.	Ja

Code	Vragen	Toelichting	Antwoord
L5	Voldoen de teksten die de leerlingen lezen aan de tekstkenmerken inhoud (thema), vorm (taalgebruik en structuur) en bedoeling (doel), zoals gespecificeerd in Checklist 1 van deze regeling en zoals nader omschreven in het Referentiekader?	De structuur van de teksten is eenvoudig en duidelijk. De 1F teksten hebben een lagere informatiedichtheid dan de 2F teksten. Ook sluiten de teksten aan bij de doelgroep. De teksten voldoen aan de tekstkenmerken zoals beschreven in de checklist. Ook voldoen de teksten aan de tekstkenmerken in het Referentiekader.	Ja

Code	Vragen	Toelichting	Antwoord
L6.1	Wordt in de toets als kenmerk van de taakuitvoering begrijpen getoetst?	In tabel 7 van de wetenschappelijke verantwoording (pagina 21) wordt een overzicht gegeven van de verdeling van de verschillende taakuitvoeringen. Begrijpen wordt getoetst.	Ja
L6.2	Wordt in de toets als kenmerk van de taakuitvoering interpreteren getoetst?	In tabel 7 van de wetenschappelijke verantwoording (pagina 21) wordt een overzicht gegeven van de verdeling van de verschillende taakuitvoeringen. Interpreteren wordt getoetst.	Ja
L6.3	Optioneel: Wordt in de toets als kenmerk van de taakuitvoering evalueren getoetst (conform de Toetswijzer PO is het onderwerp 'evalueren' een optioneel en niet verplicht onderdeel van de toets)?	Het zesde kenmerk, evalueren, is volgens het Beoordelingskader optioneel. Dit kenmerk is niet opgenomen in de DOE. (verantwoording pag. 19)	n.v.t.
L6.4	Worden in de toets de verschillende kenmerken van de taakuitvoering evenwichtig getoetst binnen het subdomein fictionele, narratieve en literaire teksten?	Kenmerken van de taakuitvoering zijn: 1. techniek en woordenschat, 2. begrijpen, 3. interpreteren, 4. samenvatten, 5. Opzoeken. Bij zakelijke teksten worden alle kenmerken (1 t/m 5) getoetst en bij fictionele, narratieve en literaire teksten kenmerk 2 en 3. (zie verantwoording pagina 19 en tabel 7, pagina 21). De verschillende taakuitvoeringen worden evenwichtig getoetst en er is gekozen voor afwisseling in de taakuitvoeringen.	Ja

#### 3.4.2 Taalverzorging



Code	Vragen	Toelichting	Antwoord
T1.1	Worden de leerlingen getoetst op regels voor lettergreepgrenzen?	Zie wetenschappelijke verantwoording tabel 9, 10, 11 en 12 op pagina 25 en 26. Lettergreepgrenzen wordt getoetst.	Ja
T1.2	Worden de leerlingen getoetst op morfologische spelling?	Zie wetenschappelijke verantwoording tabel 9, 10, 11 en 12 op pagina 25 en 26. Morfologische spelling wordt getoetst.	Ja
T1.3	Worden de leerlingen getoetst op regels voor werkwoordspelling?	Zie wetenschappelijke verantwoording tabel 9, 10, 11 en 12 op pagina 25 en 26. Werkwoordspelling wordt getoetst.	Ja
T1.4	Worden de leerlingen getoetst op overige regels?	Zie wetenschappelijke verantwoording tabel 9, 10, 11 en 12 op pagina 25 en 26. Overige spellingsregels worden getoetst.	Ja
T1.5	Worden de leerlingen getoetst op leestekens?	Zie wetenschappelijke verantwoording tabel 9, 10, 11 en 12 op pagina 25 en 26. Interpunctie wordt getoetst.	Ja
T1.6	Verantwoordt de toetsaanbieder dat er een evenwichtige verdeling is tussen opgaven rond niet-werkwoordspelling, werkwoordspelling en interpunctie?	In tabel 9 (pag. 25) van de wetenschappelijke verantwoording staat een inhoudelijke verdeling, per toetspad. Voor taalverzorging is de verdeling duidelijk. In het Excel-document staat elke vraag, inclusief kenmerk/label. De kenmerken en labels zijn passend.	Ja
T1.6a	In geval van een digitale adaptieve toets (CAT of MST), verantwoordt de toetsaanbieder op basis van welke beslisregel (het algoritme in het geval van een CAT en het module-design in geval van een MST) er wordt gegarandeerd dat er een evenwichtige verdeling is tussen opgaven rond niet-werkwoordspelling, werkwoordspelling en interpunctie?	De aanbieder legt het design uit aan de hand van fig. 1 en fig. 2 (pag. 12 en 13). Per toetspad wordt het minimum/maximum van opgaven per categorie gegeven. De verdeling is in elk toetspad evenwichtig.	Ja

### 3.5 Terrein rekenen

#### 3.5.1 Verdeling toetsopgaven over domeinen in toetsmatrijs en toetssamenstelling

Code	Vragen	Toelichting	Antwoord
D.G	Ligt het percentage toetsopgaven uit het domein Getallen tussen de 30% en 40% van het totaal aantal toetsopgaven van het terrein rekenen?	Het percentage opgaven uit het domein Getallen is 33%. Zie wetenschappelijke verantwoording 3.7.4, tabel 13.	Ja
D.Vh	Ligt het percentage toetsopgaven uit het domein Verhoudingen tussen de 20% en 30% van het totaal aantal toetsopgaven van het terrein rekenen?	Het percentage opgaven uit het domein Verhoudingen is 25%. Zie wetenschappelijke verantwoording 3.7.4, tabel 13.	Ja
D.M/mk	Ligt het percentage toetsopgaven uit het domein Meten & meetkunde tussen de 20% en 30% van het totaal aantal toetsopgaven van het terrein rekenen?	Het percentage opgaven uit het domein Meten & Meetkunde is 25%. Zie wetenschappelijke verantwoording 3.7.4, tabel 13.	Ja
D.Vb	Ligt het percentage toetsopgaven uit het domein Verbanden tussen de 15% en 20% van het totaal aantal toetsopgaven van het terrein rekenen?	Het percentage opgaven uit het domein Verbanden is 17%. Zie wetenschappelijke verantwoording 3.7.4, tabel 13.	Ja

#### 3.5.2 Verdeling toetsopgaven over onderdelen

Code	Vragen	Toelichting	Antwoord
O.A	Is iedere toetsopgave, met uitzondering van onderdeel A, onder één onderdeel geplaatst?	In tabel 14, prg 3.7.4 van de wetenschappelijke verantwoording wordt een verdeling van de onderdelen per toetspad beschreven. Iedere toetsopgave, met uitzondering van onderdeel A, is onder één onderdeel geplaatst.	Ja
O.B	Is het percentage eigen toetsopgaven uit onderdeel B afgerond ten minste 20 procent van het totaal aantal toetsopgaven van het terrein rekenen?	Per toetspad varieert dit van 23% tot 38%. In de lineaire toets is het 29%	Ja
O.C	Is het percentage toetsopgaven uit onderdeel C afgerond ten minste 30 procent van het totaal aantal toetsopgaven van het terrein rekenen?	Per toetspad varieert dit van 65% tot 75%. In de lineaire toets is het 71%.	Ja
O.BC	Is het percentage toetsopgaven uit onderdeel C hoger dan het percentage toetsopgaven uit onderdeel B?	Zie O.B en O.C. In alle toetspaden en in de lineaire toets is het percentage uit onderdeel C hoger dan het percentage uit onderdeel B. Dit geldt ook voor de lineaire toets.	Ja

#### 3.5.3 Opgaven met en zonder context

Code	Vragen	Toelichting	Antwoord
C1	Is het percentage toetsopgaven van het type contextopgaven voor het domein Getallen ten minste 30 procent van het totaal aantal toetsopgaven van het domein Getallen?	In prg 3.7.4 van de wetenschappelijke verantwoording geeft de aanbieder in tabel 15 een overzicht van de verdeling van opgaven met en zonder context. Dit varieert per toetspad van 50% tot 62%. Voor de lineaire toets is het	Ja

Code	Vragen	Toelichting	Antwoord
		56%.	
C2	Is het percentage toetsopgaven van het type contextloze opgaven voor het domein Getallen ten minste 20 procent van het totaal aantal toetsopgaven van het domein Getallen?	Dit varieert per toetspad van 38% tot 50%. Voor de lineaire toets is het 44%.	Ja
C3	Is het percentage toetsopgaven van het type contextopgaven voor het domein Verhoudingen ten minste 30 procent van het totaal aantal toetsopgaven van het domein Verhoudingen?	Dit varieert per toetspad van 67% tot 75%. Voor de lineaire toets is het 75%.	Ja
C4	Is het percentage toetsopgaven van het type contextloze opgaven voor het domein Verhoudingen ten minste 20 procent van het totaal aantal toetsopgaven van het domein Verhoudingen?	Dit varieert per toetspad van 25% tot 33%. Voor de lineaire toets is het 25%.	Ja

### 3.5.4 Gebruik kladpapier

Code	Vragen	Toelichting	Antwoord
K1	Mogen de leerlingen bij het beantwoorden van de (digitale) toetsopgaven fysiek en/of niet program-meerbaar kladpapier gebruiken?	Leerlingen mogen kladpapier gebruiken. Zie de wetenschappelijke verantwoording prg. 3.7.2 pagina 27.	Ja
K2	Wordt in de handleiding voor leraren aangegeven dat zij leerlingen nadrukkelijk dienen te wijzen op het toegestane gebruik van fysiek en/of niet programmeerbaar kladpapier?	Leerkrachten worden gewezen op het gebruik van kladpapier. Daarnaast staan in de instructieschermen voor elk reken toetsblok tips voor het gebruik van het kladpapier.	Ja

### 3.5.5 Gebruik rekenmachine

Code	Vragen	Toelichting	Antwoord
R1	Mogen de leerlingen bij het beantwoorden van de (digitale) toetsopgaven een niet-grafische en/of niet-programmeerbare rekenmachine gebruiken conform het in de Toetswijzer PO voorgeschreven aantal toetsopgaven van de doorstroomtoets?	Het gebruik van een rekenmachine is niet toegestaan. In de toelichting op deze kwaliteitseis staat een minimum van 0% van het in de toetswijzer beschreven percentage opgaven waarbij een rekenmachine mag worden gebruikt voor de grondbewerkingen (optellen, aftrekken, delen en vermenigvuldigen). De DOE overheidsdoorstroomtoets voldoet aan dit minimum.	Ja

### 3.6 Toetsopgaven Nederlandse taal en rekenen op niveau 1F en 2F/1S

Code	Vragen	Toelichting	Antwoord
TN1	Zijn de opgaven van de toets voor de terreinen Nederlandse taal en rekenen breed samengesteld qua referentieniveaus 1F en 2F / 1S?	De tabellen 3, 4 en 5 (wetenschappelijke verantwoording prg 3.4, pagina 15 tot en met 17) maken inzichtelijk hoe de toetsen zijn samengesteld qua referentieniveaus. De samenstelling is passend voor de verschillende toetspaden. De makkelijkere toetspaden voor het onderdeel rekenen bevatten meer opgaven op 1F niveau en de moeilijkere paden bevatten meer 1S. Voor Lezen en Taalverzorging is de samenstelling tussen de verschillende toetspaden gelijk. De lineaire toets is breed samengesteld en bevat voldoende opgaven voor 1F en 2F/1S.	Ja

## 4 Organisatorische aspecten

### 4.1 Leerlingrapport

Code	Vragen	Toelichting	Antwoord
LR1	Stelt de toetsaanbieder voor elke individuele leerling die de doorstroomtoets aflegt een leerlingrapport op?	Zie document leeswijzer rapportages DOE24. Voor iedere leerling wordt een rapport opgesteld.	Ja
LR2	Bevat het leerlingrapport het toetsadvies, uitgedrukt in één van de toetsadviescategorieën?	Zie document leeswijzer rapportages DOE24. Het rapport bevat het toetsadvies. De aanbieder hanteert de door de rijksoverheid vastgestelde toetsadviescategorieën.	Ja
LR3	Geeft het leerlingrapport aan welk referentieniveau Lezen de leerling beheerst?	Zie document leeswijzer rapportages DOE24.	Ja
LR4	Geeft het leerlingrapport aan welk referentieniveau Taalverzorging de leerling beheerst?	Zie document leeswijzer rapportages DOE24.	Ja
LR5	Geeft het leerlingrapport aan welk referentieniveau Rekenen de leerling beheerst?	Zie document leeswijzer rapportages DOE24.	Ja

Code	Vragen	Toelichting	Antwoord
LR6	Bevat het leerlingrapport een toelichting bij het toetsadvies, de beheersing van de referentieniveaus en eventuele optionele getoetste onderdelen van de leerling die (ook) voor ouders, verzorgers, voogden en afnemers begrijpelijk is?	Het document 'leeswijzer rapportages DOE24.PDF' geeft uitleg over de leerlingrapportage, uitgebreide leerlingrapportage en groepsrapportage. De leerlingrapportage is geschreven voor leerlingen, de andere twee rapportages voor afnemers (leerkrachten). De aanbieder heeft toegelicht dat de informatie in de leerlingrapportage ook voor ouders/verzorgers/voogden bedoeld is. Er worden in de leerlingrapportage zo min mogelijk vaktermen gebruikt. De leerlingrapportage is helder en begrijpelijk.	Ja

## 4.2 Leerlingen

Code	Vragen	Toelichting	Antwoord
LL1	Kunnen alle leerlingen, ook met specifieke ondersteuningsbehoeften (uitgezonderd leerlingen die vallen onder de ontheffingsgronden), deelnemen aan de toets?	Zie LL1.a t/m/ LL1.e. Er worden verschillende toetsversies aangeboden, zoals een Braille-versie, een DoS/TOS-versie en een aangepaste DOE overheidsdoorstrooftoets. Hierdoor is de DOE overheidsdoorstrooftoets ook toegankelijk voor leerlingen met specifieke ondersteuningsbehoeften.	Ja
LL1.a	Is de toets aantoonbaar geschikt, al dan niet door extra ondersteuning, voor leerlingen met een visuele beperking?	Er zijn verschillende mogelijkheden voor leerlingen met een visuele beperking. Er is een digitaal vergrootglas beschikbaar als hulpmiddel. In de reguliere toets is rekening gehouden met de twee meest voorkomende vormen van kleurenblindheid. Verder is er een aangepaste DOE-toets beschikbaar voor leerlingen met een andere vorm van kleurenblindheid. Voor blinde leerlingen wordt een Braille versie aangeboden.	Ja
LL1.b	Is de toets aantoonbaar geschikt, al dan niet door extra ondersteuning, voor dove of slechthorende leerlingen of leerlingen met een taalontwikkelingsstoornis?	Er is een DoS-versie beschikbaar voor deze leerlingen. In deze versie zijn opgaven met geluidsfragmenten (spelling dictee) vervangen door opgaven zonder geluidsfragmenten.	Ja
LL1.c	Is de toets aantoonbaar geschikt, al dan niet door extra ondersteuning, voor leerlingen met dyslexie?	Er wordt synthetische verklanking en vergroting aangeboden. In de wetenschappelijke verantwoording prg. 2.4 pagina 8 wordt verder beschreven hoe de DOE overheidsdoorstrooftoets geschikt is gemaakt voor leerlingen met dyslexie.	Ja
LL1.d	Is de toets aantoonbaar geschikt, al dan niet door extra ondersteuning, voor leerlingen met dyscalculie?	Er wordt tijdverlenging aangeboden.	Ja
LL1.e	Is de toets aantoonbaar geschikt, al dan niet door extra ondersteuning, voor leerlingen met andere beperkingen dan genoemd in LL1.a t/m LL1.d?	Er is geen aparte versie voor leerlingen met aandachtstoornis of stoornis in het autistisch spectrum. In de wetenschappelijke verantwoording prg. 2.4 pagina 8 beschrijft de aanbieder hoe er bij de constructie rekening is gehouden met deze leerlingen. Leerlingen met een taalontwikkelingsstoornis kunnen de aangepaste versie DOE DoS/TOS maken.	Ja
LL1.f	Wordt beschreven wat de mogelijkheden zijn als de geboden ondersteuning niet toereikend is?	Mocht een van de aangeboden versie niet toereikend zijn, dan is contact met de klantenservice mogelijk.	Ja
LL2	Kunnen alle leerlingen gebruikmaken van (de in Toetswijzer PO voorgeschreven) hulpmiddelen?	Alle leerlingen kunnen gebruik maken van kladpapier. Verder wordt een overzicht gegeven van de beschikbare hulpmiddelen, zoals vergroting en voorleesfunctie (Zie wetenschappelijke onderbouwing pagina 8).	Ja

## 4.3 Afname doorstrooftoets

Code	Vragen	Toelichting	Antwoord
A1	Is er een mogelijkheid voor een tweede afnamemoment van een inhoudelijk en psychometrische vergelijkbare variant van de in de eerste ronde afgenomen doorstrooftoets (voor leerlingen die de eerste afname zijn verhinderd)?	In het geval van een tweede afnamemoment wordt de reguliere DOE toets ingezet. Volgens de aanbieder is de geheimhouding van de DOE in de reguliere afnameperiode voldoende gewaarborgd om dezelfde toets in te kunnen zetten.	Ja
A2.1	Staat in het toetsreglement/de verantwoordingsdocumenten vermeld hoe de directeur de leerlingen voor de toets moet aanmelden?	In het stappenplan voor aanmelden staat beschreven hoe een school zich kan aanmelden.	Ja
A2.2	Staat in het toetsreglement/de verantwoordingsdocumenten vermeld welke hulpmiddelen leerlingen mogen gebruiken?	Op pagina 8 van de wetenschappelijke verantwoording wordt beschreven welke hulpmiddelen gebruikt mogen worden. Voor de DOE overheidsdoorstrooftoets kunnen alle leerlingen de volgende hulpmiddelen gebruiken: Vergrootglas, markeerstift, kladpapier, voorleesfunctie en voor de dicteopgaven een letterplankje.	Ja
A2.3	Worden in het toetsreglement/de verantwoordingsdocumenten de voorwaarden aan de (wijze van) geheimhouding van de toetsopgaven door zowel de toetsaanbieder als de deelnemende scholen beschreven?	Voorwaarden worden beschreven in het toetsreglement paragraaf 1.5 en zijn duidelijk.	Ja
A2.4	Worden in het toetsreglement/de verantwoordingsdocumenten de voorwaarden aan de (wijze van) waarop toezicht op de leerlingen wordt gehouden beschreven?	In paragraaf 1.7 van het toetsreglement worden de instructies voor tijdens de afname helder aangegeven	Ja
A2.5	Worden in het toetsreglement/de verantwoordingsdocumenten de eisen beschreven waar de ICT omgeving (computerconfiguratie) op de school aan moet voldoen om de digitale toets af te kunnen nemen?	De eisen voor de ICT omgeving worden duidelijk aangegeven in het toetsreglement, hoofdstuk 2.	Ja

## 4.4 Beveiligingsaspecten

Code	Vragen	Toelichting	Antwoord
RB1	Heeft de toetsaanbieder een toetsveiligheidsplan opgesteld met daarin opgenomen een risicoanalyse en een PDCA cyclus voor het beschermen van de beschikbaarheid, integriteit en vertrouwelijkheid van informatie (BIV)?	De aanbieder is ISO-27001 gecertificeerd. Deze kwaliteitseis is door de certificering niet van toepassing.	n.v.t.
RB1.1	Zijn de functies en verantwoordelijkheden van de functionarissen van de toetsaanbieder die betrokken zijn bij de inhoud en de beveiliging van de doorstroomtoets beschreven?	De aanbieder is ISO-27001 gecertificeerd. Deze kwaliteitseis is door de certificering niet van toepassing.	n.v.t.
RB1.2	Hebben de functionarissen van de toetsaanbieder die betrokken zijn bij de inhoud en de beveiliging van de doorstroomtoets aantoonbaar een geheimhoudingsverklaring ondertekend?	De aanbieder is ISO-27001 gecertificeerd. Deze kwaliteitseis is door de certificering niet van toepassing.	n.v.t.
RB1.3	Beschermt de toetsaanbieder het toetsconstructieproces door middel van een PDCA cyclus van beveiligingsprocedures, inclusief aantoonbare maatregelen ter voorkoming van het voortijdig uitlekken van de toetsinhoud?	De aanbieder is ISO-27001 gecertificeerd. Deze kwaliteitseis is door de certificering niet van toepassing.	n.v.t.
RB1.4	Is voor het toetsafnameproces een PDCA cyclus van beveiligingsmaatregelen geïmplementeerd, inclusief aantoonbare maatregelen ter voorkoming van het voortijdig uitlekken van de toetsinhoud?	De aanbieder is ISO-27001 gecertificeerd. Deze kwaliteitseis is door de certificering niet van toepassing.	n.v.t.
RB1.5	Zijn de fysieke en elektronische locaties waar de toetsaanbieder examens, systemen en materialen bewaart, aantoonbaar voldoende beveiligd?	De aanbieder is ISO-27001 gecertificeerd. Deze kwaliteitseis is door de certificering niet van toepassing.	n.v.t.
RB1.6	Is de digitale en fysieke informatie omtrent de examenketen zowel tijdens opslag als verzending aantoonbaar voldoende beveiligd?	De aanbieder is ISO-27001 gecertificeerd. Deze kwaliteitseis is door de certificering niet van toepassing.	n.v.t.
RB1.7	Verzorgt de toetsaanbieder periodiek training over beveiligingsbewustheid aan haar functionarissen die betrokken zijn bij de inhoud en de beveiliging van de doorstroomtoets?	De aanbieder is ISO-27001 gecertificeerd. Deze kwaliteitseis is door de certificering niet van toepassing.	n.v.t.
RB2	Waarborgt de toetsaanbieder, met behulp van passende technische en organisatorische beveiligings- maatregelen, de privacy en persoonsgegevens van de afnemer, leerling en ouder conform de Algemene Verordening Gegevensbescherming (AVG)?	De aanbieder is ISO-27001 gecertificeerd. Deze kwaliteitseis is door de certificering niet van toepassing.	n.v.t.
RB3	Heeft de toetsaanbieder een gedegen en compleet actieplan voor het omgaan met diefstal van (delen van) de toetsinhoud en voor het doen verwijderen van internetpagina's die (delen van) de toetsinhoud onthullen?	De aanbieder is ISO-27001 gecertificeerd. Deze kwaliteitseis is door de certificering niet van toepassing.	n.v.t.
RB4	Heeft de toetsaanbieder een gedegen en compleet fraudepreventieplan opgesteld met daarin opgenomen een PDCA cyclus voor de beveiliging van het toetsproces en de toetsinhoud?	De aanbieder is ISO-27001 gecertificeerd. Deze kwaliteitseis is door de certificering niet van toepassing.	n.v.t.
RB4.1	Zijn er gedegen en complete procedures voor het ontdekken en evalueren van verdachte toetsresultaten en wijzigingen in kenmerken van toetsonderdelen en toetsscores in de loop van de tijd?	De aanbieder is ISO-27001 gecertificeerd. Deze kwaliteitseis is door de certificering niet van toepassing.	n.v.t.
RB4.2	Controleert de toetsaanbieder regelmatig het internet en andere media op het bekend raken van (delen van) de inhoud van de toets of andere niet-openbare informatie over de doorstroomtoets?	De aanbieder is ISO-27001 gecertificeerd. Deze kwaliteitseis is door de certificering niet van toepassing.	n.v.t.
RB5	Voorziet de toetsaanbieder in een terugvaloptie, die in geval van kleine incidenten op leerling-, klas- of schoolniveau volgens een bijbehorend calamiteitenplan kan worden ingezet?	In geval van een (kleinschalig) incident kan een papieren versie van de landelijke calamiteitentoets worden ingezet. In het document Calamiteitenplan DOE24 wordt het stappenplan rondom de besluitvorming en de inzet van de noodtoets toegelicht.	Ja

## 5 Psychometrische aspecten

### 5.2 Pretest procedure

#### 5.2.1 Steekproefkader en samenstelling steekproef

Code	Vragen	Toelichting	Antwoord
N1.1	Is de gebruikte steekproef van voldoende kwaliteit, gezien het doel van de pretest?	Zie N1.1a en N1.1b	Ja
N1.1a	Is de steekproef groot genoeg?	Ja, er is rekening gehouden met een steekproef van 400 leerlingen, het vereiste aantal voor een geheel nieuwe toets (een aantal van >200 was voldoende geweest omdat, ondanks dat het een geheel nieuwe toets is, gebruik is gemaakt van de ACET opgavenbank in een andere afnamevorm). Bij de digitale pretesten kregen leerlingen willekeurig een van de varianten toegewezen waardoor er geen clustering van varianten op school of klasniveau plaatsvond. Hierdoor is er geen rekening gehouden met de ICC. De meeste opgaven zijn bij veel meer leerlingen afgenomen dan het vereiste aantal van 400 leerlingen. Eén opgave voldoet met 137 observaties niet aan de vereiste steekproefgrootte, maar is inhoudelijk relevant. Vandaar dat de aanbieder ervoor heeft gekozen om het op te nemen. De aanbieder legt uit dat de opgave alleen wordt meegenomen in de berekening van het	Ja

Code	Vragen	Toelichting	Antwoord
		toetsadvies indien het tijdens de operationele afname, net als in de pretest, goede psychometrische eigenschappen laat zien én er meer dan 1000 observaties zijn.	
N1.1b	Is de steekproef relevant?	Alle onder N1.1b genoemde variabelen zijn op minimaal 2 niveaus (i.e., 3-5 niveaus) aanwezig in de steekproef. De steekproeftrekking is beschreven evenals de omstandigheden waaronder de afname plaatsvond. De pretest is gevalideerd aan de hand van data van reguliere basisscholen die enkel de ACET afnamen (om een mogelijke bias door individuele keuzes voor adaptief toetsen uit te sluiten). De toetspopulatie en pretestpopulatie konden niet worden vergeleken omdat het gaat om een nieuwe toets. Daarom is de pretestpopulatie vergeleken met de landelijke populatie, en in geval van discrepanties met de Centrale Eindtoetspopulatie. Op belangrijke kenmerken, namelijk regio, urbanisatiegraad, schoolgrootte en schoolweging kwamen de steekproefscholen in grote lijnen overeen met de landelijke populatie. Regio was iets afwijkend (relatief veel scholen in het noorden) van landelijke populatie, maar wel vergelijkbaar met de ACET populatie. De proeftoets is slechts bij een beperkt aantal s(b)o scholen afgenomen. Hiervoor is een vereenvoudigde versie gebruikt. De gegevens van s(b)o scholen zijn niet meegenomen bij het bepalen van de itemparameters. Omdat de DOE zo ontwikkeld wordt dat zoveel mogelijk leerlingen de toets zonder problemen kunnen maken, worden bij de afname van de DOE relatief meer S(B)O-afnames verwacht dan bij de proeftoetsen en de uitgaven van de CET. Het exacte aantal S(B)O-leerlingen dat de DOE zal maken is pas na afname in 2024 bekend.	Ja
N1.2	Heeft de pretest plaatsgevonden conform het intern dataverzamelingsdesign of het gecombineerd dataverzamelingsdesign; en is het dataverzamelingsdesign van de pretest adequaat?	Bij de pretest opgaven is sprake van Scenario 1, doel A: de data van de pretest wordt enkel gebruikt voor een basis-kwaliteitscontrole ten behoeve van toetsamenstelling.	Ja
N1.3	Is, in het geval van open opgaven in een pretest, de beoordelaarsovereenstemming gegeven?	Alle opgaven worden automatisch gescoord.	n.v.t.

### 5.2.2 Kalibratie en kwaliteit van items

Code	Vragen	Toelichting	Antwoord
KA1	Zijn de itemparameters van de nieuwe items geschat met een bij de data passend psychometrisch model?	De itemparameters zijn geschat met het 2PL-model met behulp van het R-pakket dexterMML. Er wordt beargumenteerd welke opgaven toegelaten worden tot de operationele afname. Zo worden opgaven die extreem goed/slecht gemaakt worden (en waarvan de locatie-parameters niet nauwkeurig geschat konden worden) en opgaven die negatief discrimineren verwijderd uit de analyse en daarmee ook geëxcludeerd uit de DOE toets. De moeilijkheidsgraad van opgaven is voor zowel de populatie als geheel als voor de beoogde doelgroep (gezien de paden) geschat en de beoogde gemiddelde p-waarde bij de beoogde doelgroep van .60 wordt benaderd. RiR waardes zijn minimaal .10, met een gemiddelde van .30 - .37 voor de verschillende domeinen. Rit waardes zijn allen > .15. De kwaliteit van de afleiders is goed, gezien de a-waarden en de blokken waarin de opgaven worden aangeboden. Er is geen sprake van beoordeling van open vragen door externe beoordelaars en er is dus geen beoordelaarsbetrouwbaarheid berekend. Bij twee opgaven worden de beoogde p- en a-waarde niet behaald voor de populatie als geheel, maar wordt wel beargumenteerd wat de verwachte (voldoende) waarden zijn voor de beoogde doelgroep van het toetspad waarin de opgave wordt aangeboden.	Ja
KA2	Is er op gedegen en complete wijze beschreven dat de itemparameters van de geselecteerde gekalibreerde items toepasbaar zijn?	Op basis van Mantel-Haenszel DIF analyses is geconcludeerd dat er geen sprake van DIF op opgave-niveau met betrekking tot regio, urbanisatiegraad, schoolgrootte en schoolweging. Slechts voor één opgave in het domein Lezen bestaat een verwaarloosbare DIF voor geslacht. Vandaar dat de parameters toepasbaar zijn onafhankelijk van de genoemde achtergrond variabelen. DIF tussen boekjes en/of pretesten is niet onderzocht, aangezien dit pas bij de afname relevant is. De opgaven zullen bij afname in andere samenstelling voorkomen dan in de pretest en pretest gegevens worden genegeerd bij de berekening van het toetsadvies.	Ja
KA3	Worden beslisregels geëxpliciteerd om dekking van het inhoudelijke domein te waarborgen per individuele afname?	Er wordt duidelijk gemaakt hoe de toets start, hoe de keuze voor een volgend toetsblok wordt gemaakt, en wanneer de toets wordt beëindigd (na blok 3). De DOE is een MST (optioneel lineair) en kent 64 mogelijke toetspaden, welke allen de inhoudelijke domeinen dekken: Elk toetspad voor taal en lezen bevat minimaal 1 opgave uit ieder domeinen. De dekking over toetspaden wordt weergegeven in Tabel 7 (Lezen), Tabel 9 (Taalverzorging), Tabel 10 (Spelling niet-werkwoorden), Tabel 11 (Spelling Werkwoorden), Tabel 12 (Interpunctie). Tabel 13 (domeinen rekenen), Tabel 14 (onderdelen rekenen) en Tabel 15 (reken opgaven met/zonder context) laten de dekkingsgraad (%) over de toetspaden bij rekenen zien, waarbij alle percentages liggen binnen de gestelde eisen.	Ja
KA4	Is er een schatting gemaakt van de populatieparameters?	Er zijn een aantal aannames gedaan, aangezien dit een nieuwe toets is. Hierdoor is de algemene verwachting dat de DOE voornamelijk gemaakt zal worden door (een deel van de) scholen die voorheen de ACET afnamen. De verwachte populatieparameters zijn daarom gebaseerd op de ACET 2022.	Ja

Code	Vragen	Toelichting	Antwoord
KA5	Zijn de metingen van de individuele leerlingen voldoende betrouwbaar om uitspraken op individueel niveau te kunnen doen?	Aan de hand van gesimuleerde data (op basis van populatieparameters zoals beoordeeld bij KA4 en de latente correlaties tussen vaardigheden op onderdelen uit Tabel 29) is met berekeningen aangetoond dat de betrouwbaarheid van de toetscores en de lokale betrouwbaarheid rond de cesuren van de referentieniveaus en toetsadviescategorieën hoog is. De verdeling van standaardmeetfouten voor de verschillende vaardigheidsscores wordt getoond in Figuur 14, de standaardmeetfouten van de vaardigheidsscores (rondom de cesuurpunten) zijn vergelijkbaar met de standaarddeviaties van de ware vaardigheidsscores in de gesimuleerde populatie. De lokale betrouwbaarheid rondom de cesuurpunt voor de verschillende onderdelen varieert van .88 to .90 en is dus ruim boven de voorgeschreven .70. De betrouwbaarheid van de toetscore is .97 en daarmee ruim boven .80.	Ja
KA6	Is er aannemelijk gemaakt dat het gebruikte model past bij de data?	Voor elk onderdeel van de toets vond een iteratieve kalibratie plaats waarbij de digitale pretestgegevens werden gecombineerd met de data van de ACET 2022-afname. Telkens werd een 2PL-model geschat, waarbij de parameters van de gezamenlijke opgaven werden gefixeerd op de landelijk vastgestelde waarden. Als bepaalde opgaven niet goed pasten binnen het model, werden er aanpassingen gemaakt aan de toets- en itemspecificaties. Na elke aanpassing werd opnieuw een 2PL-model geschat. Het wordt ook benoemd hoe met opgaven wordt omgegaan die globaal extreem goed of slecht worden gedaan en met opgaven met grote misfit statistieken. Ook was er geen aanwijzing voor DIF op opgave niveau (zie KA2). Er wordt echter niet expliciet aangegeven dat en hoe aan verschillende voorwaarden van KA6 wordt voldaan, vooral de voorwaarde van unidimensionaliteit. Voor een inschatting van de dimensionaliteit van de DOE gebruikt de aanbieder data van de ACET 2022, omdat voor de DOE zelf nog geen data is afgenomen. De latente correlaties tussen twee subdomeinen binnen één onderdeel laten zien dat de subdomeinen voor elk onderdeel sterk samenhangen - met een minimale correlatie van .81 binnen Taalverzorging - en de data dus voldoende unidimensionaliteit toont.	Ja
KA7	Bij gebruik van een 2PL IRT model, is er een indicatie gegeven van de lokale betrouwbaarheid van de nieuwe items?	Er is variatie in gemiddelde discriminatieparameters, met gemiddelden tussen domeinen variërend van 0.99 tot 1.33, en een minimum variërend van .16 tot .42. Alle nieuwe opgaven hebben een geschatte alpha > .15, hoger dan 10% van de gemiddelde alpha van het betreffende onderdeel.	Ja

### 5.2.1 Toetsamenstelling

Code	Vragen	Toelichting	Antwoord
T1	Bevat de doorstroomtoets voor de terreinen Nederlandse taal en rekenen in ieder geval vragen met een moeilijkheidsgraad tussen de cesuur van referentieniveau 1F en 2F/1S?	Tabellen 3-5 tonen de aantallen opgaven per referentieniveau. Alle leerlingen maken voldoende opgaven van beide referentieniveaus, moeilijker 1F opgaven zijn geplaatst in moeilijke blokken en eenvoudiger 1S of 2F opgaven in de makkelijke blokken. De aanbieder stelt daarom dat "bij alle onderdelen er items zijn opgenomen met locatielparameters tussen de 1F- en 2F/1S cesuren in". De informatiecurves in Figuur 11 t/m 13 ondersteunen dit, aangezien hieruit blijkt dat de informatiecurves van de bewust toetspaden pieken bij de referentieniveaus van de beoogde doelgroep.	Ja
T2	Bevat de doorstroomtoets het door het CvTE voorgeschreven minimaal aantal gezamenlijke anker-items voor de wettelijk verplichte terreinen Nederlandse taal en rekenen? Is de volledige subset van gezamenlijke ankeritems opgenomen in (de varianten van) de doorstroom- toets? Zijn de gezamenlijke ankeritems op de in de Regeling Beoordelingsnormen Doorstroomtoetsen po beschreven wijze opgenomen in de doorstroomtoets?	De voorgeschreven opgaven van het gezamenlijk anker zijn opgenomen. De 20 ankeritems voor ieder toetsonderdeel zijn verspreid over de drie toetsblokken van het onderdeel (8 in blok 1, 4 in blok 2, 8 in blok 3). De opgaven zijn minimaal aangepast om ze niet te laten opvallen. In iedere variant van de blokken komen ankeritems voor, maar de aantallen verschillen iets over de varianten op Blok 1 (Bijv alle of niet alle opgaven in een blok). Doordat toewijzing aan blokken willekeurig plaatsvindt zullen de leerlingen die alle opgaven maken representatief zijn voor de populatie. Bij blok 2 en 3 worden in alle varianten alle ankeritems gemaakt, dit met het oog op het mogelijk moeten laten vallen van blokken bij beperkte inschrijving voor DOE.	Ja
T3	Bestaat een papieren toets of een lineaire digitale toets jaarlijks vrijwel uit nieuwe items voor de wettelijk verplichte terreinen Nederlandse taal (domeinen Lezen en Taalverzorging) en rekenen, met uitzondering van de items van het gezamenlijk anker?	De aanbieder beschrijft dat DOE een nieuwe toets en primair adaptieve toets is. Indien een school kiest voor een lineaire afname, dan maakt een leerling het gemiddelde toetspad (blokken A1-C1-F1). Het wordt gehandhaafd dat opgaven die na 1 jaar tot en met 25% exposure kennen en door meer dan 2000 leerlingen zijn beantwoord, na 1 jaar worden vernieuwd. Opgaven die tot maximaal 25% exposure per jaar hebben, worden na maximaal 3 inzetten vernieuwd. Dit betekent dat opgaven uiterlijk drie jaar worden gebruikt, maar vervangen worden zodra ze in een jaar door meer dan 25% van de leerlingen in dat jaar zijn beantwoord.	n.v.t.
T4	Heeft de itembank van een CAT een omvang van minimaal 6x het aantal items van een maximale operationele toetsversie, waarbij jaarlijks tussen de 20% en 30% van de items wordt ververs?	De DOE is (optioneel) adaptief maar een MST a.d.h.v. blokken i.p.v. een CAT. Verversing is pas relevant bij de volgende afname (2025)	n.v.t.

Code	Vragen	Toelichting	Antwoord
T4.1	Is er, in het geval van een MST die bestaat uit meerdere lagen van modules (c.q. routes), sprake van een adequate verversingsstrategie a. voor niet geselecteerde items? b. voor items met discrepanties in parameters in het kalibratiedesign en de voorlaatste operationele fase? c. voor items die uitgelekt lijken? d. voor items die over jaargangen heen anders functioneren?	De DOE is (optioneel) een MST, maar verversing is pas relevant bij de volgende afname (2025).	n.v.t.

### 5.3 Rapportage achteraf van de kwaliteit en het functioneren van de doorstroomtoets

Code	Aan te leveren tabellen	Toelichting	Antwoord
Tabel 1	Een tabel met aantallen opgaven per vaardigheid: aantal wel en aantal niet meegenomen voor de scorerapportage naar de leerlingen wegens (1) een inhoudelijke fout, (2) psychometrisch disfunctioneren, (3) DIF, (4) te weinig waarnemingen. Wanneer er zaai-items ingezet worden kunnen deze in een aparte tabel weergegeven worden. Het leveren van een lijst welk item dit betreft wordt aangeraden.	De DOE overheidsdoorstroomtoets is een nieuwe toets. Er zijn geen afnamegegevens van 2023 beschikbaar. Criterium 5.3 is niet van toepassing.	n.v.t.
Tabel 2	Een overzicht (tabel) van de percentages referentieniveaus van de wettelijk verplichte en optionele domeinen van het terrein Nederlandse taal (domein lezen en taalverzorging) en het terrein rekenen plus een overzicht van de percentages toetsadviezen (de zes categorieën) van dit jaar ten opzichte van het vorige jaar.	De DOE overheidsdoorstroomtoets is een nieuwe toets. Er zijn geen afnamegegevens van 2023 beschikbaar. Criterium 5.3 is niet van toepassing.	n.v.t.
Tabel 3	Een overzicht (tabel) van de relatie tussen de referentieniveaus en de achtergrondvariabelen van de doelgroep.	De DOE overheidsdoorstroomtoets is een nieuwe toets. Er zijn geen afnamegegevens van 2023 beschikbaar. Criterium 5.3 is niet van toepassing.	n.v.t.
Tabel 4	Een overzicht (tabel) van de relatie tussen de toetsadviezen en de achtergrondvariabelen van de doelgroep.	De DOE overheidsdoorstroomtoets is een nieuwe toets. Er zijn geen afnamegegevens van 2023 beschikbaar. Criterium 5.3 is niet van toepassing.	n.v.t.
Tabel 5	Voor wat betreft de eerder afgenomen doorstroomtoets, een overzicht (tabel) van de verdeling van de populatievaardigheid (gemiddelde en standaardafwijking) van in ieder geval de reguliere po-populatie inclusief de schattingsfout (standard error) op de vaardigheidsverdeling, een en ander op dezelfde schaal als de itemparameters.	De DOE overheidsdoorstroomtoets is een nieuwe toets. Er zijn geen afnamegegevens van 2023 beschikbaar. Criterium 5.3 is niet van toepassing.	n.v.t.
Tabel 6	Bij eerder afgenomen opgaven (in eerder jaar of bij pretest): DIF-analyse-gegevens met betrekking tot parameters van de meest recente operationele afname en de eerdere afname(n), een en ander ter onderbouwing van de bij de kwaliteitseisen T4 en T4.1 beoordeelde verversingsstrategie van de aanbieder.	De DOE overheidsdoorstroomtoets is een nieuwe toets. Er zijn geen afnamegegevens van 2023 beschikbaar. Criterium 5.3 is niet van toepassing.	n.v.t.
Tabel 7	Een overzicht (tabel) van de overeenkomsten tussen de toetsadviezen en de eerder gegeven (voorlopige) schooladviezen, bestaande uit een overzicht van het percentage consistente en niet consistente classificaties, inclusief de globale classificatiefout (c.q. het percentage misclassificaties als functie van de mate van overeenstemming tussen het voorlopig schooladvies en het toetsadvies)	De DOE overheidsdoorstroomtoets is een nieuwe toets. Er zijn geen afnamegegevens van 2023 beschikbaar. Criterium 5.3 is niet van toepassing.	n.v.t.

**Bureau van het CvTE**

Muntstraat 7  
3512 ET Utrecht  
Postbus 315  
3500 AH Utrecht  
Nederland  
www.cvte.nl

# Advies DOE

**Omschrijving** Een onderzoek van de Adviescommissie toetsen po naar de zorgvuldigheid waarmee het advies van Stichting Cito tot stand is gekomen inzake de aanvraag tot erkenning van de DOE overheidsdoorstroomtoets 2024.

**Bijlagen** Advies Stichting Cito

**Datum**

24 augustus 2023

**Onze referentie**

23.00945

## 1 AANVRAAG

Op 31-05-2023 heeft het College voor Toetsen en Examen (CvTE) een aanvraag voor de erkenning van de DOE overheidsdoorstroomtoets ontvangen en op 01-06-2023 in behandeling genomen.

## 2 BEOORDELING

Bij de beoordeling van de doorstroomtoets beoordeelt het CvTE of de toets voldoet aan de wettelijke vereisten, zoals omschreven in de Regeling beoordelingskader doorstroomtoets po. Hierbij wordt een onderscheid gemaakt tussen de onderwijskundige, organisatorische en de psychometrische aspecten van de toets. Het CvTE maakt voor het erkennen van een doorstroomtoets gebruik van de adviseurs van Stichting Cito. Stichting Cito past voor zijn advies de Regeling beoordelingskader doorstroomtoets po toe. De Adviescommissie toetsen po heeft dit oordeel bij de behandeling van de aanvraag meegewogen.

### 2.1 ONDERWIJSKUNDIGE ASPECTEN

Het advies van Stichting Cito is voor dit onderdeel zorgvuldig tot stand gekomen en de Adviescommissie toetsen po neemt het advies op dit punt over.

### 2.2 ORGANISATORISCHE ASPECTEN

Het advies van Stichting Cito is voor dit onderdeel zorgvuldig tot stand gekomen en de Adviescommissie toetsen po neemt het advies op dit punt over.

### 2.3 PSYCHOMETRISCHE ASPECTEN

Het advies van Stichting Cito is voor dit onderdeel zorgvuldig tot stand gekomen en de Adviescommissie toetsen po neemt het advies op dit punt over.



### 3 OORDEEL

De Adviescommissie toetsen po heeft geen aanwijzingen om aan te nemen dat het advies niet zorgvuldig tot stand is gekomen en adviseert het CvTE om het advies over te nemen.

Namens de Adviescommissie toetsen po,

(voorzitter)