



> Retouradres Postbus 315 3500 AH Utrecht

Bureau van het CvTE

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

Contactgegevens

030 - 28 40 700
info@cvte.nl

Datum 2 februari 2024
Betreft Besluit erkenning leerlingvolgsysteem 'Boom LVS Rekenen Wiskunde Contextrijk'

Onze referentie

CvTE-24.00420

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 of het leerlingvolgsysteem Boom LVS Rekenen Wiskunde Contextrijk voldoet aan de voorwaarden op basis van de wet. Uw verzoek is ontvangen op 16 oktober 2023 (door ons geregistreerd met correspondentienummer: 39847).

1 BESLUIT EN WETTELIJK KADER

Gelet op artikel 3a, vierde lid, Wet College voor toetsen en examens, artikel 45b, eerste lid, Wet op het primair onderwijs, heeft het college besloten het leerlingvolgsysteem Boom LVS Rekenen Wiskunde Contextrijk te erkennen voor tien jaar met ingang van de dagtekening van dit besluit 2 februari 2024.

2 MOTIVERING

Bij de beoordeling van het leerlingvolgsysteem Boom LVS Rekenen Wiskunde Contextrijk van Boom Test Onderwijs heeft het college onderzocht of de toets voldoet aan de wettelijke vereisten en de Regeling beoordelingskader voor de toetsen behorende tot leerling en onderwijsvolgsystemen PO. Het college heeft hierbij gebruik gemaakt van het advies van Stichting Cito. De Stichting heeft bij haar advisering de Regeling beoordelingskader voor de toetsen behorende tot leerling en onderwijsvolgsystemen 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 ervan 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 zal in de Staatscourant worden gepubliceerd. De adviezen worden samen met het besluit op de website van het CvTE (www.cvtetoetsenpo.nl) gepubliceerd.

Bureau CvTE

Datum

2 februari 2024

Onze referentie

CvTE CvTE-24.00420

Hoogachtend,

Het College voor Toetsen en Examens,
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

Definitief advies LVS

Stichting Cito – 22-12-2023

Boom Rekenen Wiskunde Contextrijk

Ter beoordeling van Boom LVS Rekenen Wiskunde Contextrijk geeft Stichting Cito advies aan het College voor Toetsen en Examens. In dit rapport staat het definitief advies van Stichting Cito aan het CvTE beschreven. Het definitief advies is tot stand gekomen aan de hand van het [beoordelingskader LVS-instrumenten](#). Bij dit advies zit een bijlage met daarin een overzicht van de beoordeling per kwaliteitseis (bijlage 1).

De aanvraag betreft een: Initiële indiening

LVS-instrument type: III

Doelgroep: groep 6 t/m 8

Onderdeel: rekenen/wiskunde Contextrijk

Is het beoordelingsteam tijdens de beoordeling tegen problemen aangelopen?

Ja

Toelichting indien ja: In de verantwoording op kwaliteitseis O.V.1 miste het beoordelingsteam per toets en per opgave een specificatie van het referentieniveau. Zonder deze specificatie kon niet worden geverifieerd of de verhouding 1F/1S opgaven passend is bij de doelgroep van het instrument. In de verantwoording van kwaliteitseis P.VII.3 miste het beoordelingsteam onderzoek naar de divergente validiteit. Bij kwaliteitseis P.VI.1 ontbrak in de aangeleverde documenten informatie over onderzoek naar de lokale betrouwbaarheid. Tijdens het hoor- en wederhoorgesprek zijn deze punten besproken. De aanbieder heeft op basis van het gesprek aanvullende informatie aangeleverd, die vervolgens is beoordeeld.

1. Aanvraag en proces

Op 18 oktober 2023 heeft Stichting Cito de documenten voor de beoordeling van het Boom LVS Contextrijk (CR) ontvangen. Op 18 oktober zijn de documenten gedeeld met twee beoordelaars, een psychometrisch beoordelaar en een onderwijskundig beoordelaar. Op 16 november is er overleg geweest tussen de beoordelaars en de editors van Stichting Cito (het beoordelingsteam). De editors van Stichting Cito hebben de gezamenlijke bevindingen verwerkt tot een conceptadvies. Het conceptadvies is op 24 november 2023 gedeeld met het CvTE ter voorbereiding op het hoor- wederhoorgesprek op 29 november 2023.

2. Algemene informatie

De algemene informatie over het leerlingvolgsysteem komt uit het document gebruikershandleiding en verantwoording Boom LVS RW Contextrijk.

Boom LVS RW Contextrijk is onderdeel van het leerlingvolgsysteem van Boom test onderwijs. Het Boom LVS RW Contextrijk is een nieuw LVS-instrument als aanvulling op de bestaande Boom LVS toetsen.

Doelgroep en meetpretentie

Boom LVS RW Contextrijk is bestemd voor leerlingen in groep 6 tot en met 8 van het Nederlandse basisonderwijs.

De toets meet de rekenvaardigheid van leerlingen op het gebied van contextrijk rekenen. Leerlingen wordt gevraagd hun rekeninzicht, kennis en vaardigheden toe te passen in contextopgaven die doorgaans bestaan uit situaties die in het dagelijks leven kunnen voorkomen én betekenisvol of voorstelbaar voor leerlingen zijn. Het LVS bestaat uit 3 Blokken met in totaal 6 toetsen. Het gebruiksdoel van Boom LVS Contextrijk is driedelig: Boom LVS CR kan worden ingezet voor niveaubepaling van de leerlingen, zowel individueel als per groep. Het LVS kan daarnaast worden ingezet voor de voortgangsbepaling van de leerlingen. Door verschillende meetmomenten kunnen leerlingen met zichzelf vergeleken worden. Het Boom LVS RW Contextrijk maakt voor deze vergelijking gebruik van de vaardigheidsscore en het didactische leeftijdsequivalent (DLE). Ten derde kan het LVS CR helpen bij het signaleren van achterstand in rekenvaardigheid van het oplossen van contextopgaven.

Inhoudelijke verantwoording

De toets is opgebouwd uit drie blokken, waarbij ieder blok de leerstof weerspiegelt van één leerjaar in het basisonderwijs (vanaf groep 6). De Boom LVS-toets Contextrijk bestaat merendeels uit open opgaven, maar kent ook gesloten (meerkeuze) opgaven. De drie blokken Boom LVS Contextrijk hebben elk twee versies (versie A en versie B). De versies zijn parallelle versies. Dit betekent dat beide versies uit dezelfde soort opgaven zijn samengesteld, die ook in dezelfde volgorde worden aangeboden. De opgaven doen een beroep op dezelfde rekenoperaties die leerlingen moeten uitvoeren om de opgaven op te lossen. De aanbieder geeft aan dat het Boom LVS CR aansluit bij de kerndoelen voor het primair onderwijs en bij de referentieniveaus.

Overzicht van de aangeleverde documenten:

De aanbieder heeft de volgende documentatie aangeleverd:

- Leeswijzer Boom LVS RW leeswijzer Contextrijk
- Ouderbrochure (pdf)
- Gebruikershandleiding en verantwoording Boom LVS Contextrijk
- Instructiekaart Boom Contextrijk
- Toetsboekje (pdf) + antwoordenboek (pdf)
- Scoreblad
- Boom LVS RW CR Normen (Excel)(pdf)
- Boom LVS RW CR Bijlage vraagonzuiverheid(pdf)
- ISO-certificaat

3. Kwaliteitscriteria beoordelingskader

Boom LVS CR is een nieuw instrument en valt binnen de Regeling beoordelingskader voor de toetsen behorende tot leerling- en onderwijsvolgsystemen PO onder LVS instrumenten Type III. Het instrument is beoordeeld op de volgende kwaliteitseisen:

Aspecten	Paragraaf	Kwaliteitseis
Onderwijskundige aspecten	2.3	O.III, O.IV en O.VI
Psychometrische aspecten	3.3	P.IV t/m P.VIII
Organisatorische aspecten	4.1	B.I

Hieronder volgt een kort overzicht van de belangrijkste bevindingen van het beoordelingsteam op de drie verschillende aspecten. Een overzicht van de beoordeling per kwaliteitseis is te vinden in de bijlage.

Onderwijskundige aspecten

De toetsen behorende bij het Boom LVS CR bevatten zowel open vragen als ook meerkeuze opgaven. De opgaven voldoen aan de in de checklist voor het beoordelen van de kwaliteit van observatie-categorieën en toetsopgaven genoemde richtlijnen. Om de contexten te beschrijven wordt in de toetsen taal/tekst gebruikt. Hierdoor wordt een bepaald leesvaardigheidsniveau van leerlingen verwacht. De zinnen zijn kort gehouden en het taalgebruik is eenduidig. Het beoordelingsteam is van mening dat de gebruikte tekst past bij de doelgroep van het instrument en dat in Boom LVS CR de taal geen belemmering vormt bij het meten van de rekenvaardigheid op het gebied van contextrijk rekenen. Dit wordt bevestigd in de verantwoording op kwaliteitseis P.VII.1, waarin de unidimensionaliteit van het instrument wordt verantwoord. Door het gebruik van herkenbare contexten zijn de opgaven relevant voor de doelgroep. De A-versie en B-versie van de toetsen zijn gelijk qua opbouw, domeindekking en moeilijkheid. De aanbieder geeft aan dat het Boom LVS CR aansluit bij het referentiekader en rapporteert over het behaalde referentieniveau. Om te beoordelen of een uitspraak over het behaalde referentieniveau voldoende onderbouwd kan worden, miste het beoordelingsteam aanvankelijk een overzicht van referentieniveau per opgave. De verhouding 1F/1S was niet gespecificeerd, waardoor het voor het beoordelingsteam onvoldoende inzichtelijk is of de verhouding binnen de toetsen op orde is. In het hoor- en

wederhoorgesprek is aan de aanbieder gevraagd specificatie van de referentieniveaus aan te leveren.

Psychometrische aspecten

De steekproef van Boom LVS CR is duidelijk beschreven en omvat voldoende observaties om de modelparameters te schatten. Ook het dataverzamelingsdesign wordt door het beoordelingsteam als voldoende beschouwd. De drie blokken van opgaven zijn voldoende met elkaar gelinkt en de unidimensionaliteit is door middel van PCA onderzocht. De passing van het Rasch-model is goed, met uitzondering voor Blok 6, die ook op andere gebieden een uitschieter lijkt te zijn. De afwijkingen in Blok 6 en de grafische weergave van de passing van het model op itemniveau zijn met de aanbieder tijdens het hoor-en wederhoor gesprek besproken.

Twee eisen waren in de initiële beoordeling door het beoordelingsteam als onvoldoende beschouwd: het onderzoek naar de lokale betrouwbaarheid (P.V1.1) en informatie over de divergente validiteit (P.VII). Dit kwam tijdens het hoor- wederhoorgesprek ter sprake en de aanbieder heeft hierover aanvullende informatie aangeleverd.

Organisatorische aspecten

De aanbieder is ISO 27001 gecertificeerd met een geldigheid tot 01-11-2025. De kwaliteitscriteria in paragraaf 4.1 zijn door de ISO-certificering voldoende verantwoord.

4. Hoor- en wederhoorgesprek

Op woensdag 29 november 2023 heeft onder regie van het CvTE het hoor-wederhoor gesprek plaatsgevonden. Tijdens het gesprek zijn de in het conceptadvies aangegeven punten besproken: de verhouding van de referentieniveaus, het onderzoek naar de lokale betrouwbaarheid bij specifieke punten op de vaardigheidsschaal en het onderzoek naar de divergente validiteit. De aanbieder heeft tijdens het gesprek de gevraagde informatie gepresenteerd. Per toetsblok zijn de referentieniveaus getoond en is het onderzoek naar de lokale betrouwbaarheid gepresenteerd. Een onderzoek naar de divergente validiteit kon niet worden gepresenteerd. In plaats van onderzoek naar de divergente validiteit, heeft de aanbieder voorlopig een schatting toegevoegd die aanduidt dat de divergente validiteit voldoende zou kunnen worden aangetoond. Zodra voldoende gegevens van verschillende LVS onderdelen beschikbaar zijn, zal de aanbieder informatie over de divergente validiteit in de handleiding toevoegen. Er is afgesproken dat alle gepresenteerde informatie schriftelijk wordt aangeleverd bij het CvTE. De behandeltermijn is tot het moment van de aanlevering opgeschort.

5. Proces richting definitief advies

Op 7 december heeft de aanbieder de aanvullende informatie op de in het hoor- en wederhoorgesprek besproken punten aangeleverd. De informatie is op 7 december gedeeld met het beoordelingsteam. In de aanvullende informatie geeft de aanbieder een overzicht van de referentieniveaus per opgave en per toets. Op basis van dit overzicht heeft het beoordelingsteam vastgesteld, dat de opgaven op het juiste referentieniveau zijn gelabeld en dat de verhouding 1F/1S opgaven over de verschillende blokken in orde is. De aanbieder

heeft in de aanvullende informatie ook onderzoek naar de lokale betrouwbaarheid toegevoegd, i.e., de meetprecisie bij specifieke punten op de vaardigheidsschaal, en informatie over de divergente validiteit aangeleverd, waardoor de eisen P.VI.1 en P.VII door het beoordelingsteam zijn goedgekeurd. Tijdens het hoor- wederhoorgesprek kwamen ook de psychometrische afwijkingen van blok 6 ter sprake en de aanbieder heeft deze afwijkingen ook met aanvullende informatie voldoende kunnen toelichten (zie toelichting P.IV.2).

6. Oordeel en advies

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

Paragraaf	Kwaliteitseis	Codering kwaliteitseis	Voorlopig oordeel
2.3.1	Uitgangspunten van de toetsconstructie	O.III	Voldoende
2.3.2	Inhoudsvaliditeit	O.IV	Voldoende
2.3.3	Referentieniveaus	O.V	Voldoende
2.3.4	Inzicht in leervorderingen	O.VI	Voldoende
3.3.1	Kwaliteit van de dataverzameling	P.IV	Voldoende
3.3.2	Normering	P.V	Voldoende
3.3.3	Betrouwbaarheid	P.VI	Voldoende
3.3.4	Constructvaliditeit	P.VII	Voldoende
3.3.5	Volgaspect	P.VIII	Voldoende
4.1	Afname	B.I	Voldoende

Stichting Cito geeft op basis van de verstrekte verantwoordingsdocumenten een positief advies voor de initiële beoordeling van Boom LVS RW Contextrijk.

Dit advies is vastgesteld op 22 december 2023 te Arnhem.

Bijlage 1: Toelichting beoordeling per kwaliteitscriterium

2.3 Onderwijskundige aspecten

2.3.1 Uitgangspunten van de toetsconstructie

Code	Vragen	Toelichting	Antwoord
O.III.1	Is de afnamevorm van het instrument gedegen en compleet gedefinieerd?	<i>Boom LVS CR voor groep 6, 7 en 8 bestaat uit een papieren toetsboekje. Het heeft drie blokken; voor elk leerjaar één blok. Elk blok bestaat uit twee vergelijkbare toetsen (parallele versies): een A-versie en een B-versie. De toetsen van circa 40 opgaven bestaan uit open opgaven (kort antwoord vragen) en meerkeuze opgaven (Zie gebruikershandleiding §2.3, blz 20). In de instructiekaart voor leerkrachten worden heldere stappen voor de afname beschreven. De tijdsduur wordt in dit document aangegeven en er staat beschreven welke hulpmiddelen zijn toegestaan. In de instructiekaart en in de gebruikershandleiding wordt beschreven dat scholen kunnen kiezen om één variant of beide varianten van het blok af te nemen in een leerjaar.</i>	<i>Ja</i>
O.III.2	Is (zijn) de doelgroep(en) van het instrument gedegen en compleet gedefinieerd?	<i>Het Boom LVS CR is bedoeld voor groep 6-7-8 van het primair onderwijs(PO). In de gebruikershandleiding (§1.3, blz. 13) wordt aangegeven dat de toets ook ingezet kan worden in het speciaal basisonderwijs. Zie instructiekaart punt 3 en §1.1 van de gebruikershandleiding, blz 15.</i>	<i>Ja</i>
O.III.3	Is het meetdoel van het instrument gedegen en compleet gedefinieerd?	<i>De toets meet rekenvaardigheid van leerlingen op het gebied van contextopgaven. De inhoudelijke beschrijving van de toets (Hoofdstuk 2 van de gebruikershandleiding) is duidelijk en compleet toegelicht. De toetsen sluiten aan bij de kerndoelen en de referentieniveaus. De verschillende domeinen van Rekenen-wiskunde worden getoetst. Elk blok</i>	<i>Ja</i>

Code	Vragen	Toelichting	Antwoord
		<i>weerspiegelt de leerstof van dat betreffende leerjaar. Zie handleiding §1.2 en §1.3, blz. 13+14.</i>	
O.III.4	Is er een actuele gedegen en gedetailleerde inhoudelijke theoretische inkadering van het instrument?	<i>In de gebruikershandleiding §2.1 en §2.2 handleiding, blz. 17+18 staat beschreven welke kadering gebruikt is en welke documenten ten grondslag liggen aan dit instrument. De inhoudelijke theoretische inkadering is gedegen beschreven. LVS Contextrijk sluit aan op de referentieniveaus en de kerndoelen PO. Bij de ontwikkeling van het instrument is gebruik gemaakt van verschillende documenten, waaronder Concretisering referentieniveaus en Tussendoelen rekenen-wiskunde voor het primair onderwijs.</i>	Ja
O.III.5	Is het gebruiksdoel van het instrument gedegen en compleet gedefinieerd?	<i>De gebruiksdoelen staan beschreven in de gebruikershandleiding §1.3, blz. 13 en 14. Hier worden drie gebruiksdoelen aangegeven: bepalen ontwikkelingsniveau, ontwikkeling volgen en signaleren rekenachterstand. De leerlingenrapportages geven percentielscores (uitgedrukt in niveau I-V en A-E), het behaalde referentieniveau en de didactische leeftijdsequivalent (DLE-score) (zie §3.3, blz 24,25 handleiding.</i>	Ja

2.3.2 Inhoudsvaliditeit

Code	Vragen	Toelichting	Antwoord
O.IV.1	Is de toetsmatrijs een adequate, schematische representatie van het meetdoel?	<i>In de gebruikershandleiding §2.2, 2.3 en 2.4, wordt in diverse tabellen de verdeling over de domeinen getoond. Voor elk blok wordt in de gebruikershandleiding aangegeven hoe deze is samengesteld (tabellen 2.3 t/m 2.6).</i>	Ja

Code	Vragen	Toelichting	Antwoord
		<ul style="list-style-type: none"> • Circa 40 opgaven met bijbehorende punten (per opgave 1 punt) • Open of meerkeuze-opgaven met aantal vragen: tabel 2.6 • Toegestane hulpmiddelen: pen en kladpapier: instructiekaart punt 7: kladpapier, potlood en gum, geen rekenmachine • Toetsduur: instructiekaart punt 6 => 50 tot 60 minuten is indicatie, leerlingen mogen langer werken als ze dat nodig hebben. <p>De aanbieder toont per blok de verdeling over de domeinen in tabel 2.4, p.22 van de gebruikershandleiding. De verdeling van de verschillende domeinen sluit aan bij de verhouding die gebruikt wordt bij de doorstroomtoetsen PO. In Bijlage I van de gebruikershandleiding is een overzicht gegeven van de inhoudsdoelen per opgave. Het beoordelingsteam is van mening dat de verschillende inhoudsdoelen representatief terugkomen in de blokken.</p>	
O.IV.2	Voldoen de toetsopgaven in het instrument aan de kwaliteitscriteria voor toetsopgaven: relevantie, objectiviteit, efficiëntie, specificiteit en neutraliteit?	<p>De gekozen contexten zijn herkenbaar voor leerlingen en maken de opgaven relevant voor de doelgroep. In de toetsen zitten geen opgaven waarbij meerdere antwoorden verdedigbaar zijn. Verder bevat het LVS CR geen aanstootgevende contexten of opgaven waarbij voorkennis vereist is. De aanbieder erkent dat er bij het maken van contextrijke rekenopgaven een beroep wordt gedaan op de leesvaardigheid van leerlingen. In de constructie is hiermee rekening gehouden door korte en eenduidige zinnen te gebruiken om de context te beschrijven. Afbeeldingen worden functioneel gebruikt.</p>	Ja

Code	Vragen	Toelichting	Antwoord
O.IV.3	Voldoen de toetsopgaven in het instrument aan de constructievoorschriften voor gesloten vragen, open vragen en/of nieuwe vraagtypes?	<i>De opgaven voldoen aan de constructievoorschriften. De toets varieert in vraagtypes. De instructiezin/vraagzin is bij elke opgave helder, zodat een leerling weet wat er verwacht wordt. De afleiders bij de meerkeuze opgaven zijn plausibel en vaak afgeleid uit veel gemaakte denkfouten.</i>	<i>Ja</i>

2.3.3, Referentieniveaus

Code	Vragen	Toelichting	Antwoord
O.V.1	Bevat de verzameling van LVS-instrumenten van het type (III) van groep 6-8 voor wat betreft de wettelijk verplichte voorgeschreven terreinen Nederlandse taal en rekenen én de optionele domeinen van taal, de inhouden van 1F en 2F/1S?	<i>De toetsen Contextrijk bevatten voor wat betreft het wettelijk verplichte voorgeschreven terrein rekenen de inhouden van de referentieniveaus 1F en 1S. De verhouding 1F/1Sverschilt per blok. In blok 4 worden nog nauwelijks 1S opgaven aangeboden. In blok 5 is de verhouding 1F/1S opgaven gelijk en in blok 6 ligt het zwaartepunt bij 1S. De toetsen Contextrijk sluiten met deze oplopende verhouding aan bij wat er van het betreffende leerjaar verwacht kan worden.</i>	<i>Ja</i>
O.V.1a	Indien er sprake is van een CAT (adaptieve toets op itemniveau), voldoet de gehele itembank van de verzameling van LVS-instrumenten van het type (III) van groep 6-8 voor wat betreft de bij O.V.1 genoemde domeinen aan de kwaliteitseis O.V.1?	<i>Boom LVS CR is een papieren toets, geen adaptieve toets. Hierdoor is deze eis niet van toepassing.</i>	<i>n.v.t.</i>
O.V.1b	Indien er sprake is van een MST (adaptieve toets op moduleniveau), voldoet het gehele design van de verzameling van LVS-instrumenten van het type (III) van groep 6 tot en met 8 voor wat betreft de bij O.V.1	<i>Boom LVS CR is een papieren toets, geen adaptieve toets. Hierdoor is deze eis niet van toepassing.</i>	<i>n.v.t.</i>

	genoemde domeinen aan de kwaliteitseis O.V.1?		
--	---	--	--

2.3.4 Inzicht in leervorderingen

Code	Vragen	Toelichting	Antwoord
O.VI.1	Stelt de toetsaanbieder voor elke individuele leerling een leerlingrapport op?	<i>Er volgt per toets een individueel rapport waarop de niveauscore van de leerling is aangegeven. Bij de niveauscore gaat het om een relatieve score. Dit wordt uitgedrukt in de percentielscore op het rapport. (H3.4 van de gebruikershandleiding, blz.20 en 21) In de gebruikershandleiding zijn twee voorbeeldrapporten opgenomen (p.38 en 40).</i>	Ja
O.VI.2	Bevat het leerlingrapport het gemeten niveau van de leerling en geeft het rapport de leervordering van de leerling weer?	<i>De voorbeeldrapporten in de gebruikershandleiding laten zien dat het gemeten niveau wordt gerapporteerd. De leervordering wordt inzichtelijk gemaakt in een individueel overzicht. De leerkracht kan voor elke leerling een individueel overzicht aanmaken. Op pagina 42 en 44 van de gebruikershandleiding staan voorbeelden van een individueel overzicht.</i>	Ja
O.VI.3	Bevat het leerlingrapport een toelichting bij het gemeten niveau en de leervorderingen van de leerling die (ook) voor ouders, verzorgers, voogden en docenten begrijpelijk is?	<i>De rapportages bevatten moeilijk te begrijpen begrippen. De aanbieder legt in de ouderbrochure in heldere taal uit waarom leerlingen getest worden, waarop getest wordt en hoe de rapportage gelezen kan worden. Voor docenten zijn groepsrapportages beschikbaar en wordt in hoofdstuk 3 van de handleiding beschreven hoe de rapportage gelezen moet worden. De extra toelichting voor de specifieke doelgroep (ouders/voogden of leerkrachten) bij het rapport maakt de informatie begrijpelijk.</i>	Ja

Code	Vragen	Toelichting	Antwoord
O.VI.4	Is de informatie in het leerlingrapport een goede basis voor de leerkracht om de leervorderingen te evalueren en vervolgstappen te formuleren?	<i>In §3.4 van de gebruikershandleiding wordt beschreven hoe scores geïnterpreteerd moeten worden. Dit is concreet en helder uitgelegd met een aantal casussen erbij als voorbeeld. Om de leervorderingen goed te kunnen volgen moet een leerkracht de individuele overzichten gebruiken. Dit wordt duidelijk uitgelegd in de gebruikershandleiding. Leerkrachten kunnen ook beschikken over groepsrapportages, zodat ook op groepsniveau vervolgstappen geformuleerd kunnen worden.</i>	Ja

3. Psychometrische aspecten

3.3 Kwaliteitseisen psychometrische aspecten LVS-instrumenten – type (III) en (IV)

3.3.1 Kwaliteit van de dataverzameling

Code	Vragen	Toelichting	Antwoord
P.IV.1	Is de steekproef van leerlingen groot genoeg en representatief voor de landelijke populatie?	<i>De uiteindelijke steekproef bestond uit totaal 2634 leerlingen afkomstig van in totaal 41 basisscholen. Omdat Boom LVS-toets Rekenen-Wiskunde Contextrijk drie blokken van oplopend niveau kent en er aparte normeringen zijn opgesteld voor ieder blok, zijn er in feite drie verschillende steekproeven. De meeste leerlingen hebben twee blokken gemaakt, waardoor het aantal toetsgegevens hoger ligt dan het leerlingaantal. Uit Tabel 5.2 blijkt dat het totaal aantal toetsafnames voor blok 4, blok 5 en blok 6 respectievelijk 1352, 1637 en 1054 was. Dat is ruim voldoende voor het schatten van de modelparameters. De achtergrondvariabelen denominatie, schoolgrootte en stedelijkheid tonen aanvaardbaar kleine afwijkingen van het landelijk gemiddelde. Om de verdeling van de beide belangrijkste achtergrondvariabelen regio en schoolweging zo dicht mogelijk die van de populatie te laten benaderen, is er</i>	Ja

Code	Vragen	Toelichting	Antwoord
		<p><i>op deze variabelen een weging toegepast. Afwijkingen in de steekproef t.o.v. de populatie kunnen door middel van weging van de steekproef inderdaad voor een belangrijk deel worden gecorrigeerd.</i></p> <p><i>Ter controle van de schoolafhankelijkheid heeft de aanbieder de intraklassecorrelatiecoëfficiënt (ICC) berekend (Tabel 5.17, pagina 71). Over het algemeen is de ICC met <0.15 redelijk laag, met een opvallende afwijking bij DL= 43 in steekproef Blok 5 met een waarde van 0.350. Dat kan volgens de toetsaanbieder worden verklaard doordat een aantal relatief zwak-presterende groepen in de steekproef van DL= 43 zijn opgenomen.</i></p>	
P.IV.2	Indien er sprake is van een onvolledig dataverzamelingsdesign: is het dataverzamelingsdesign adequaat?	<p><i>De Boom LVS-toets Rekenen-Wiskunde Contextrijk kent drie blokken, namelijk Blok 4, 5 en 6. De voortgang van de leerling wordt door middel van de scores op deze blokken met elkaar vergelijkt op één en dezelfde onderliggende latente schaal. Hiertoe zijn de blokken met elkaar gelinkt en is een vaardigheidsschaal berekend door middel van toepassing van het Rasch-model uit de IRT (Item Response Theorie) dat de waarschijnlijkheid modelleert van een fout/goed antwoord op een opgave, gegeven de vaardigheid van een leerling. Voor de parameterschatting van dit meetmodel (i.e., de moeilijkheidsparameter) op basis van de Joint Maximum Likelihood-methode (JML) m.b.v. het programma Jmetrik (Meyer, 2014), dient er zowel verticale linking (tussen de blokken voor de verschillende leerjaren) als horizontale linking (tussen versie A en versie B van de LVS-toets) plaats te vinden.</i></p> <p><i>De unidimensionaliteit van de toetsen is door middel van PCA onderzocht en is aannemelijk. Om de passing van het Rasch-model te beoordelen, is gebruikgemaakt van de Infit- en Outfit-statistieken. Uit Tabel IV blijkt dat er geen opgaven zijn met een Infit-waarde of een Outfit-waarde < 0.50 en ></i></p>	Ja

Code	Vragen	Toelichting	Antwoord
		<p><i>2.00, en slechts enkele opgaven die niet aan de strengere criteria waarden < 0.75 en > 1.25 voldoen. Omdat het slechts een enkele opgave betreft, vindt het beoordelingsteam dit aanvaardbaar.</i></p> <p><i>Op basis van KTT is te zien dat er slechts enkele opgaven met minder goede r_{ir}-waarden zijn; de item-restcorrelaties zijn voor bijna alle opgaven voldoende. Ook kunnen de nauwkeurigheden 'c' van de parameterschattingen van de opgaven als goed worden beoordeeld (dit geldt niet voor Blok 6 Versie B wat door een ontoereikende kalibratiesteekproef zou kunnen worden verklaard).</i></p> <p><i>Ook is de passing van het Rasch-model grafisch beoordeeld door middel van de verwachte scoregrafieken. In Bijlage V wordt voor ieder blok een voorbeeld gegeven van een goede, een gemiddelde en een slechtere passing. De Infit- en Outfit-waarden worden bevestigd door de grafieken; over het algemeen liggen de geobserveerde scores dicht bij de op grond van het Rasch-model verwachte waarden. Afwijkingen hiervan kunnen meestal worden verklaard doordat het meerkeuze-opgaven betreft waarbij de gokkans een rol speelt, welke niet in het Rasch-model wordt gemodelleerd. Om de overall-fit van het Rasch-model te beoordelen zijn voor de kalibratiegegevens voor iedere toets voor de ongelinkte Rasch-parameters met het TAM R-package (Robitzsch, 2022) de effectmaat MADaQ3 berekend. Uit Tabel 4.7 blijkt dat deze MADaQ3-waarden voor de A-versies van de blokken (0.0319, 0.0309 en 0.0354 voor resp. Blok 4, 5 en 6) meer dan voldoende zijn. De B-versies laten hogere waarden zien voor Blok 4 en Blok 5 (0.0744 en 0.0726), maar nog wel aanvaardbaar. Voor de B-versie van Blok 6 is de MADaQ3-waarde van 0.1000 echter te hoog.</i></p> <p><i>Tenslotte is met behulp van Jmetrik nog nagegaan of het 2PL-model (moeilijkheidsparameter en discriminatieparameter) tot essentieel andere resultaten leidt dan het Rasch-model. Het meer complexe 2PL-model laat een goede passing van de opgaven zien (item fit statistiek S bij slechts een</i></p>	

Code	Vragen	Toelichting	Antwoord
		<p>tot twee opgaven van een blok $p < 0.05$). Tabel 4.8 toont de geschatte gemiddelde vaardigheidsscores van de twee modellen (d.m.v. EAP = Expected a Posteriori) en de correlaties tussen beide voor de blokken 4 t/m 6 van Versie A. Uit deze tabel blijkt dat de correlaties voor alle blokken hoog zijn (0.9911, 0.9937 en 0.9943 voor resp. Blok 4, 5 en 6), hetgeen betekent dat de berekende vaardigheidsscore op basis van het Rasch-model die van het 2PL-model dicht nadert.</p> <p>Om de parallelliteit van versie A en versie B te onderzoeken zijn de gemiddelden en standaarddeviaties van beide versies berekend. Conform de verwachting blijkt uit Tabel 4.9 dat de verschillen zeer klein zijn (Cohen's d is -0.025, 0.019 en -0.060 voor resp. Blok 4, 5 en 6) en tevens de correlaties tussen beide versies hoog zijn (0.841, 0.910 en 0.856 voor resp. Blok 4, 5 en 6). De aanbieder heeft ervoor gekozen om via de Stocking-Lord methode de parallelliteit van beide versies verder te onderzoeken wat ook als voldoende wordt beschouwd.</p> <p>Verschillen van Blok 6 zouden kunnen worden verklaard door een ontoereikende kalibratiesteekproef. Dit is onderbouwd door een update van de aanbieder in de aanvullende informatie met een aangevulde steekproef (+14 leerlingen) die al een verbetering laat zien. Op dit moment verzamelt de aanbieder aanvullende gegevens om deze oneffenheid in de dataverzameling verder te kunnen corrigeren.</p> <p>De aanbieder toont de passing van het Rasch-model (op itemniveau) ook grafisch.</p>	

3.3.2 Normering

Code	Vragen	Toelichting	Antwoord
Absoluut normeren			

Code	Vragen	Toelichting	Antwoord
P.V.1	Is de keuze voor de standaardbepalingsmethode gedegen gemotiveerd en op de juiste wijze uitgevoerd?	<i>Dit kwaliteitsaspect heeft betrekking op het vaststellen van absolute normen en is derhalve niet van toepassing.</i>	<i>n.v.t.</i>
P.V.2	Zijn de experts met inhoudelijke vakdeskundigheid die de standaarden (cesuren) vaststellen naar behoren geselecteerd en getraind?	<i>Dit kwaliteitsaspect heeft betrekking op het vaststellen van absolute normen en is derhalve niet van toepassing.</i>	<i>n.v.t.</i>
P.V.3	Is er voldoende overeenstemming tussen de experts?	<i>Dit kwaliteitsaspect heeft betrekking op het vaststellen van absolute normen en is derhalve niet van toepassing.</i>	<i>n.v.t.</i>
Relatief normeren			
P.V.4	Als de normeringssteekproef afwijkt van de kalibratiesteekproef, zijn de normgroepen groot genoeg?	<i>De normeringssteekproef wijkt niet af van de kalibratiesteekproef.</i>	<i>n.v.t.</i>
P.V.5	Als de normeringssteekproef afwijkt van de kalibratiesteekproef, zijn de normgroepen representatief?	<i>De normeringssteekproef wijkt niet af van de kalibratiesteekproef.</i>	<i>n.v.t.</i>
P.V.6	Zijn de normen correct bepaald?	<i>De Boom LVS-toets Rekenen-Wiskunde Contextvrij is continu genormeerd wat betekent dat de toets in principe gedurende het hele schooljaar kan worden afgenomen (m.u.v. Blok 4, de toets voor groep 6, die vanaf de maand november kan worden afgenomen). Bij de normering wordt uitgegaan van de maand waarin de toets is afgenomen, waarbij de systematiek van de didactische leeftijd (DL) in onderwijsmaanden wordt gehanteerd. Met behulp van het model van Tellegen kan de gewenste omzetting van ruwe score naar een</i>	<i>Ja</i>

Code	Vragen	Toelichting	Antwoord
		<p><i>voorspelde standaard normaal verdeelde normscore Z voor elke didactische leeftijd worden gerealiseerd.</i></p> <p><i>De empirische Z-scores worden goed voorspeld door de gehanteerde modellering en de normaliteit van de verdeling van T-scores kan worden aangenomen.</i></p> <p><i>De grafieken 1 t/m 3 voor resp. Blok 4 t/m Blok 6 laten zien dat in het algemeen wordt voldaan aan de voorwaarden dat een hogere vaardigheidsscore leidt tot een hogere normscore en dat de normen in de loop van het basisonderwijs strenger worden in de periode waarvoor het model geldt.</i></p> <p><i>Een uitzondering op de tweede voorwaarde is de grafiek van de normering van Blok 6 waar de normeringsfuncties voor DL = 59 en DL = 60 (i.e., de laatste twee onderwijsmaanden van groep 8) de andere lijnen snijden. De aanbieder verklaart dit doordat een deel van de leerlingen licht terugvallen in hun niveau van contextrijke rekenvaardigheid na het maken van de Eindtoets.</i></p>	

3.3.3 Betrouwbaarheid

Code	Vragen	Toelichting	Antwoord
P.VI.1	Zijn of worden de betrouwbaarheidsgegevens van de metingen correct berekend?	<p><i>Voor het globale betrouwbaarheidsonderzoek van de Boom LVS-toets Rekenen-Wiskunde Contextrijk zijn de interne consistentie in de vorm van Cronbach's alfa, McDonald's omega en de parallelle betrouwbaarheid bepaald (beide uit de KTT).</i></p> <p><i>De aanbieder heeft naar aanleiding van het hoor-wederhoor gesprek aanvullende informatie voor de lokale betrouwbaarheid aangeleverd.</i></p> <p><i>De lokale betrouwbaarheid is aan de uiteinden van de schaal in het algemeen lager, wat betekent dat het moeilijker is om de vaardigheid nauwkeurig te bepalen wanneer een leerling een blok heel goed of</i></p>	<i>Ja</i>

Code	Vragen	Toelichting	Antwoord
		<i>heel slecht maakt. Het wordt dan door de aanbieder aanbevolen, indien het mogelijk is, terug- dan wel vooruit te toetsen (de leerling een blok laten maken dat meer past bij zijn of haar niveau).</i>	
P.VI.2	Zijn de betrouwbaarheidsgegevens voldoende gezien de conclusies en eventuele beslissingen die met het instrument genomen worden?	<i>Tabel 6.1 en Tabel 6.2 presenteren per DL en blok de betrouwbaarheden Cronbach's alpha en McDonald's omega voor resp. versie A en versie B. Beide betrouwbaarheden liggen steeds dicht in de buurt van 0.90 of hoger. De parallelle betrouwbaarheden worden in Tabel 6.3 getoond als de Pearson correlatiecoëfficiënten tussen Versie A en Versie B. Uit deze tabel blijkt dat de parallelle betrouwbaarheden hoog zijn (0.841, 0.910 en 0.856 voor resp. Blok 4, Blok 5 en Blok 6) en slechts een fractie lager dan de eerder gepresenteerde betrouwbaarheden Cronbach's alpha en McDonald's omega.</i>	Ja
P.VI.3	Is er een indicatie gegeven van het percentage misclassificaties van de verschillende categorieën van het LVS-instrument van het type (III) of (IV)?	<i>Bijlage VI toont voor Versie A de accuratesse en de accuratesse plus-minus 1 van de niveau-indeling A-E en I-V voor de verschillende blokken en de verschillende normeringsmomenten, welke ook in Tabel 6.4 worden weergegeven. Uit de tabellen in Bijlage VI en Tabel 6.4 blijkt dat de accuratessewaarden variëren van 0.68 - 0.74. Een accuratesse van 0.65 is bij vier cesuren een acceptabele waarde en deze grens wordt gehaald. Voor de accuratesse plus-minus 1 geldt een richtlijn van 0.95. De tabellen in Bijlage VI en Tabel 6.4 tonen dat aan deze richtlijn wordt voldaan, omdat de accuratesse plus-minus 1 in de buurt van 0.99 ligt. Verder tonen de tabellen in Bijlage VI ook nog aan dat het heel erg weinig voorkomt dat een leerling twee niveaus lager of hoger wordt ingedeeld en dat de accuratesse in de hoogste en laagste niveau-indeling het grootste is.</i>	Ja

3.3.4 Constructvaliditeit

Code	Vragen	Toelichting	Antwoord
P.VII.1	Is aangetoond dat er sprake is van unidimensionaliteit?	<i>In paragraaf 4.3.3 (Eendimensionaliteit) is met behulp van ongeroteerde PCA (Principale Component Analyse) met tetrachorische correlatie unidimensionaliteit van de Boom LVS-toets CR aannemelijk gemaakt. Op basis van deze bevinding (en het onderzoek naar de inhoudsvaliditeit) mag er worden uitgegaan van unidimensionaliteit voor het geassumeerde Rasch-model als meetmodel en dat de opgaven één latente trek of construct, Contextrijk, meten (en dus niet onbedoeld mede beïnvloed wordt door een andere vaardigheid, zoals bijv. taalvaardigheid).</i>	<i>Ja</i>
P.VII.2	Is aangetoond dat de itemkwaliteit op orde is?	<i>Uit een tweede pilotstudie van november 2021 bleken de opgaven over het algemeen door de leerlingen moeilijk gevonden. Gemiddeld wordt ongeveer de helft van de opgaven correct gemaakt met de gemiddelde p-waarde iets onder de 0.50 (0.48, 0.47 en 0.47 voor resp. Blok 4, Blok 5 en Blok 6; zie Tabel 4.2). Op basis van deze gegevens en de door de leerling gegeven antwoorden is een aantal opgaven in de blokken vereenvoudigd door de getaloperatie die de opgave vereist, te vereenvoudigen. Tevens is een aantal opgaven omgezet naar een meerkeuzeopgave. Uit Tabel 4.2 is ook af te leiden dat de gemiddelde item-rest correlaties met waarden van 0.30, 0.33 en 0.42 voor resp. Blok 4 t/m Blok 6 voldoende zijn. Afgezien van de relatief lage p-waarden kan uit de resultaten van de tweede pilotstudie geconcludeerd worden dat de opgaven goede psychometrische eigenschappen hebben en de itemkwaliteit op orde is. Een aantal opgaven met lage item-rest correlatie (zie de vetgedrukte opgaven met item-restcorrelaties < 0.20 in de tabellen</i>	<i>Ja</i>

Code	Vragen	Toelichting	Antwoord
		<i>van Bijlage IV) en slechte passing van het Rasch-model zijn tekstueel aangepast en, indien nodig, vereenvoudigd.</i>	
P.VII.3	Is er onderzoek uitgevoerd naar de convergente en divergente validiteit?	<p><i>De aanbieder heeft onderzoek uitgevoerd naar convergente validiteit (i.e., toetsen die hetzelfde construct meten moeten hoog correleren). Dit is gedaan door de Boom LVS-toets Rekenen-Wiskunde Contextrijk te relateren aan de Referentiesets Rekenen-Wiskunde 1F en 1S, die door Stichting Cito in opdracht van het CvTE zijn ontwikkeld. Tabel 7.3 laat zien dat de Cronbach's alpha betrouwbaarheden van de referentiesets goed zijn (0.870 en 0.835 voor resp. het 1F-niveau en het 1S-niveau).</i></p> <p><i>Door de Boom LVS-toets Rekenen-Wiskunde Contextvrij met deze ankersets van het CvTE te verankeren, kan eerst een cesuur worden bepaald waarmee kan worden vastgesteld of een leerling het fundamentele 1F-niveau dan wel het 1S-streefniveau heeft gehaald. Voor de cesuurbepaling van het 1S-niveau zijn alleen leerlingen uit groep 8 betrokken, terwijl bij de cesuurbepaling van het 1F-niveau ook leerlingen uit groep 7 zijn meegenomen.</i></p> <p><i>Vervolgens zijn de correlaties tussen de referentiesets en de relevante blokken van de Boom LVS-toets Rekenen-Wiskunde Contextvrij berekend (zie Tabel 7.4). Uit Tabel 7.4 blijkt dat deze correlaties voldoen aan de verwachting van relatief hoge correlaties, hetgeen dus opgevat kan worden als een indicatie voor convergente validiteit (>0.72).</i></p> <p><i>De aanbieder kan op dit moment geen expliciete informatie over de divergente validiteit aanleveren, omdat deze ten tijde van de</i></p>	<i>Ja</i>

Code	Vragen	Toelichting	Antwoord
		<p><i>ontwikkeling van het instrument nog niet beschikbaar was. De aanbieder geeft aan deze informatie in de handleiding op te nemen zodra deze beschikbaar is (zodra voldoende scholen deze toets en ook het Boom LVS Technisch Lezen en het LVS Begrijpend Lezen hebben afgenomen). Op basis van een schatting geeft de aanbieder aan te verwachten dat de correlaties tussen de twee andere LVS-instrumenten en Boom Contextrijk lager uitvallen dan 0.5. Op basis van deze toelichting en toezegging heeft het beoordelingsteam deze eis als voldoende beoordeeld.</i></p> <p><i>Bij de tussentijdse check van dit instrument zal het onderzoek naar de divergente validiteit beoordeeld worden.</i></p>	
P.VII.4	Is aangetoond dat er geen sprake is van item bias?	<p><i>In het document 'Bijlage Onderzoek Vraagzuiverheid' wordt op pag. 2 (1^e alinea) vermeld dat voor het onderzoek naar item bias (vraagzuiverheid) is gekeken naar stedelijkheid (stedelijkheid versus platteland) en schoolgewicht (hoog en laag). Tabel 1 laat zien dat de referentiegroep voor stedelijkheid en schoolgewicht resp. platteland en laag is, terwijl de focusgroep voor stedelijkheid en schoolgewicht resp. stedelijkheid en hoog is.</i></p> <p><i>Om vraagzuiverheid te detecteren in de normeringssteekproef is gebruikgemaakt van de aangepaste methodiek van Angoff's delta methode (Magis & Facon, 2012) uit het R-package DifR. In de aangepaste methodiek van Angoff's delta methode geldt als criterium van vraagzuiverheid dat de zogenaamde 'distance' tussen -1.5 en 1.5 moet liggen (zie de tabellen t/m 12 uit de bijlage Vraagzuiverheid). Uit Tabel 2 en Tabel 3 blijkt dat er voor Versie A voor zowel Blok 4, Blok 5 als voor Blok 6 geen vraagzuiverheid is</i></p>	Ja

Code	Vragen	Toelichting	Antwoord
		<p>gevonden voor <i>stedelijkheid</i> en <i>schoolgewicht</i>. Voor <i>Blok 4</i> van <i>Versie B</i> is er <i>één</i> opgave voor <i>stedelijkheid</i> en zijn er twee opgaven voor <i>schoolgewicht</i> gedetecteerd (<i>distance</i> vetgedrukt in de tabellen 1 t/m 12 uit de <i>Bijlage</i>).</p> <p><i>Tabel 2</i> en <i>Tabel 3</i> tonen ook dat er bij <i>Blok 6</i> van <i>Versie B</i> een aantal <i>vraagonzuivere</i> opgaven zijn gedetecteerd voor zowel <i>stedelijkheid</i> als <i>schoolgewicht</i>. De aanbieder is voornemens om in het schooljaar 2023-2024 hiervoor aanvullende gegevens te verzamelen teneinde deze vraag te kunnen beantwoorden. De relatief hoge <i>c</i>-constanten van de parameterschattingen van de opgaven van <i>Blok 6</i> <i>Versie B</i> en de te hoge <i>globale</i> <i>passingsmaat</i> <i>MADaQ3</i> van het <i>Rasch</i>-model van <i>Blok 6</i> <i>Versie B</i> zijn meer dan waarschijnlijk ook te herleiden tot de relatief kleine <i>steekproefomvang</i> van <i>Blok 6</i> <i>Versie B</i> en deze aanvullende gegevens kunnen dus ook hiervoor gebruikt worden (zie noot 7, pag. 55) van de <i>gebruikerhandleiding</i> en <i>verantwoording</i>).</p>	
P.VII.5	Is onderzoek uitgevoerd naar de prestaties van relevante subgroepen?	<p>Er is onderzoek uitgevoerd naar verschillen voor <i>schoolgewicht</i>. In <i>Tabel 7.1</i> worden de scores voor de vijf verschillende <i>niveaugroepen</i> van <i>schoolweging</i> gepresenteerd.</p> <p>Uit <i>Tabel 7.1</i> blijkt dat er sprake is van een effect van de <i>schoolweging</i> op de <i>gemiddelde</i> <i>vaardigheidsscore</i>. De effectmaten η^2 laten een matig effect ($\eta^2 \geq 0.06$) zien (0.09, 0.11 en 0.05 voor resp. <i>Blok 4</i>, <i>Blok 5</i> en <i>Blok 6</i>) voor de <i>gemiddelde</i> <i>vaardigheidsscore</i>. De verschillen in <i>vaardigheidsscores</i> tussen de leerlingen op een school met het laagste (< 23) en met het hoogste (≥ 37) <i>schoolgewicht</i> zijn ongeveer 3, 4 en 4 voor resp. <i>Blok 4</i>, <i>Blok 5</i> en <i>Blok 6</i>.</p>	Ja

3.3.5 Volgaspect

Code	Vragen	Toelichting	Antwoord
P.VIII.1	Is er een voldoende empirische onderbouwing van de schaal waarop de groei van een leerling wordt uitgedrukt? Wordt groei op een adequate manier gemeten?	<p><i>Uit de resultaten van het kalibratieonderzoek blijkt dat de opgaven van Boom LVS Contextrijk door middel van het gehanteerde Rasch-model op een unidimensionele onderliggende vaardigheidsschaal afgebeeld kunnen worden. Dit laatste maakt het mogelijk om zowel het niveau als de groei/voortgang van de contextrijke rekenvaardigheid te kunnen volgen op één en dezelfde doorlopende vaardigheidsschaal gedurende het gehele basisonderwijs. Het niveau en de groei van een leerling wordt via grafieken weergegeven, zie bijvoorbeeld Grafiek 3.1, waarbij op de horizontale as de onderwijsmaand (= DL) en op de verticale as de vaardigheidsscore staat.</i></p> <p><i>In de groeicurve van grafiek 3.1 is een zwarte lijn weergegeven. Dit is de 50-percentielijn, oftewel de gemiddelde ontwikkeling. Daarnaast zijn op basis van de normeringssteekproef ook de 5%-, de 35%-, de 65%- en de 95%-lijnen berekend. Op basis hiervan worden vijf gebieden (niveaus) in de groeicurve weergegeven (zie Tabel 3.5). Als de vaardigheidsscores van een leerling consequent onder het gearceerde gebied liggen (niveau 1), kan dat een signaal voor een probleem zijn. Indien de vaardigheidsscores steeds boven het gearceerde gebied liggen (niveau 5), gaat het om een excellente leerling op het gebied van contextrijke rekenvaardigheid. In paragraaf 3.8 van de gebruikershandleiding en verantwoording Boom LVS RW Contextrijk worden enkele casusbeschrijvingen gegeven waarin de aanbieder uitlegt hoe de groei van leerlingen wordt weergegeven.</i></p>	Ja

Code	Vragen	Toelichting	Antwoord
P.VIII.2	Wordt de betrouwbaarheid van de groei op die schaal adequaat weergegeven?	<i>De (ruwe) toetscores van de leerling zijn op een (gekalibreerde) eendimensionale vaardigheidsschaal (Rasch-schaal) geplaatst aan de hand waarvan de groei van een leerling via de groeigrafiek gevolgd kan worden. Omdat gebruik wordt gemaakt van IRT, kan de betrouwbaarheid van de groei op de Rasch-schaal adequaat weergegeven worden. Voor een interpretatie van de vaardigheidsscore kan de gemiddelde groei van de vaardigheidsscore behulpzaam zijn. Om te bepalen of er sprake is van werkelijke groei worden 90%-betrouwbaarheidsintervallen gebruikt. Daarnaast kan ook gekeken worden naar de percentielscores, de gebruikelijke twee niveau-indelingen (I-V en A-E) en de DLE-scores van opeenvolgende toetsafnames en deze te vergelijken tussen afnames om groei vast te stellen.</i>	Ja
P.VIII.3	Worden er gegevens verstrekt over hoe groei geïnterpreteerd dient te worden?	<i>In Hoofdstuk 3 van de gebruikershandleiding wordt groei onderbouwd met diverse rapportagemogelijkheden (individueel rapport en individueel overzicht). Ook wordt in Hoofdstuk 3 van de gebruikershandleiding aan de hand van een individueel leerlingrapport de interpretatie van groei voor de leerkracht op een heldere en duidelijke manier beschreven. Voor iedere normeringsmaat is op basis van het normeringsmodel de vaardigheidsscore berekend voor 50% (de 'DLE-lijn'). Vervolgens is de toename in vaardigheidsscore van normeringsmaand tot normeringsmaand berekend. In Bijlage III wordt de verwachte gemiddelde groei in vaardigheidsscore na een maand en de cumulatie van de gemiddelde groei na meerdere maanden in tabelvorm gerapporteerd. Aan de hand van voorbeelden wordt in</i>	Ja

Code	Vragen	Toelichting	Antwoord
		<p><i>paragraaf 3.3.4 toelicht hoe groei geïnterpreteerd dient te worden in deze tabellen.</i></p> <p><i>Bij het Boom LVS Contextrijk wordt een informatiebrochure voor ouders/verzorgers/voogden aangeboden, getiteld 'Boom LVS Volgen'. In die brochure wordt aan ouders/verzorgers/voogden uitgelegd hoe de contextrijke rekenvaardigheid van hun kind getoetst wordt en hoe de prestatie van hun kind in cijfers of letters uitgedrukt wordt.</i></p>	

4. Organisatorische aspecten

4.1 Afname

Code	Vragen	Toelichting	Antwoord
B.I.1	Heeft de aanbieder de functies en verantwoordelijkheden van de functionarissen die betrokken zijn bij de inhoud en de beveiliging van het LVS-instrument volledig en juist beschreven?	<i>De aanbieder is ISO 27001 gecertificeerd. De kwaliteitseisen behorende bij de organisatorische aspecten zijn hierdoor voldoende verantwoord.</i>	<i>Ja</i>
B.I.2	Hebben de functionarissen van de aanbieder die betrokken zijn bij de inhoud en de beveiliging van het LVS-instrument aantoonbaar een geheimhoudingsverklaring ondertekend?	<i>De aanbieder is ISO 27001 gecertificeerd. De kwaliteitseisen behorende bij de organisatorische aspecten zijn hierdoor voldoende verantwoord.</i>	<i>Ja</i>
B.I.3	Waarborgt de aanbieder op gedegen en complete wijze de privacy van de leerkracht, (jonge) leerling en ouder conform de Algemene Verordening Gegevensbescherming)?	<i>De aanbieder is ISO 27001 gecertificeerd. De kwaliteitseisen behorende bij de organisatorische aspecten zijn hierdoor voldoende verantwoord.</i>	<i>Ja</i>
B.I.4	Hanteert de aanbieder op gedegen en complete wijze incidentmanagement inclusief een protocol datalekken,	<i>De aanbieder is ISO 27001 gecertificeerd. De kwaliteitseisen behorende bij de organisatorische aspecten zijn hierdoor voldoende verantwoord.</i>	<i>Ja</i>

Code	Vragen	Toelichting	Antwoord
	zoals voorgeschreven door de Autoriteit Persoonsgegevens?		
B.I.5	Bevat de handleiding voor de schoolleider en leerkracht op gedegen en complete wijze informatie over de wijze waarop het instrument en de ermee verzamelde informatie op een veilige wijze gebruikt en opgeslagen dient te worden?	<i>De aanbieder is ISO 27001 gecertificeerd. De kwaliteitseisen behorende bij de organisatorische aspecten zijn hierdoor voldoende verantwoord.</i>	<i>Ja</i>
B.I.6	Controleert de aanbieder regelmatig het internet en andere media op het bekend raken van (delen) van de inhoud van het LVS-instrument of andere niet-openbare informatie over het LVS-instrument?	<i>De aanbieder is ISO 27001 gecertificeerd. De kwaliteitseisen behorende bij de organisatorische aspecten zijn hierdoor voldoende verantwoord.</i>	<i>Ja</i>



Definitief advies

'Boom LVS toets Rekenen-Wiskunde Contextrijk'

Bureau van het CvTE

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

Datum

30 januari 2024

Onze referentie

CvTE-24.00008

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 het leerlingvolgsysteem 'Boom LVS toets Rekenen-Wiskunde Contextrijk' van Boom Test Onderwijs.
Bijlagen	Advies Stichting Cito

1 AANVRAAG

Op 16 oktober 2023 heeft het College voor Toetsen en Examen (CvTE) een aanvraag voor de erkenning van het leerlingvolgsysteem 'Boom LVS toets Rekenen-Wiskunde Contextrijk' van Boom Test Onderwijs ontvangen en in behandeling genomen.

2 BEOORDELING

Bij de beoordeling van het leerlingvolgsysteem beoordeelt het CvTE of de toets voldoet aan de wettelijke vereisten, zoals omschreven in de Regeling beoordelingskader voor de toetsen behorende tot leerling- en onderwijsvolgsystemen PO (het beoordelingskader). Hierbij wordt een onderscheid gemaakt tussen de onderwijskundige en organisatorische aspecten, en de psychometrische aspecten van de toets. Het CvTE maakt voor het erkennen van een leerlingvolgsysteem gebruik van het advies van Stichting Cito. Stichting Cito past voor haar advies het beoordelingskader toe. De adviescommissie toetsen po heeft onderzocht of stichting Cito op zorgvuldige wijze het beoordelingskader heeft toegepast en tot haar advies is gekomen.

2.1 ONDERWIJSKUNDIGE ASPECTEN

Het advies van Stichting Cito is voor dit onderdeel zorgvuldig tot stand gekomen en de adviescommissie toetsen po adviseert het college het advies op dit punt over te nemen.

2.2 ORGANISATORISCHE ASPECTEN

Het advies van Stichting Cito is voor dit onderdeel zorgvuldig tot stand gekomen en de adviescommissie toetsen po adviseert het college het advies op dit punt over te nemen.

2.3 PSYCHOMETRISCHE ASPECTEN

De adviescommissie toetsen po is van oordeel dat het advies van Stichting Cito zorgvuldig tot stand is gekomen en adviseert het college het advies op dit punt over te nemen.

De adviescommissie toetsen po verzoekt het college **expliciet** om helderheid te verschaffen over de interpretatie van de punten P.V.1, P.V.2, P.V.3 en P.VII.3 uit het beoordelingskader. Voor de punten P.V.1, P.V.2 en P.V.3 dient het beoordelingskader uitgebreid te worden om ook bij het gebruik van een bestaande norm een deugdelijke linking naar deze bestaande norm te kunnen beoordelen. Bij punt P.VII.3 dient duidelijk te worden wat er onder onderzoek naar divergente validiteit wordt verstaan en wanneer deze deugdelijk wordt geacht.

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 met inachtneming van de hierboven genoemde psychometrische punten die nadere helderheid behoeven, om het advies over te nemen.

Namens de adviescommissie toetsen po,

de voorzitter