

## Definitief advies LVS Taalverzorging: Spelling Woorden en Spelling Werkwoorden

Stichting Cito, 18 december 2025

Ter beoordeling van het Boom LVS Taalverzorging: Spelling Woorden en Spelling Werkwoorden geeft Stichting Cito advies aan het College voor Toetsen en Examens. In dit rapport staat het advies van Stichting Cito voor de toetsaanbieder beschreven. Het 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: 10-jaarlijkse erkenning

LVS-instrument type: III

Doelgroep: groep 3 tot en met 8

Onderdeel: Taalverzorging

Is het beoordelingsteam tijdens de beoordeling tegen problemen aangelopen?

Ja:

Toelichting indien ja: bij de initieel ingediende informatie miste het beoordelingsteam een adequate verantwoording voor de beoordeling van kwaliteitseisen O.III.2, O.IV.1, O.IV.2, O.V.1b, P.IV.1, P.IV.2, P.V.4, P.V.5, P.VI.1, P.VI.3, P.VII.2, P.VII.4, P.VII.5 en P.VIII.2.

Tijdens het hoor- en wederhoorgesprek zijn deze kwaliteitseisen besproken. De aanbieder heeft op basis van het gesprek aanvullende informatie aangeleverd.

## 1 Aanvraag en proces

Op 10 oktober 2025 heeft Stichting Cito de documenten voor de beoordeling van het Boom leerlingvolgsysteemtoets Taalverzorging: Spelling Woorden en Spelling Werkwoorden (Boom LVS-toets Taalverzorging: Spelling) 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 13 oktober 2025 zijn de documenten gedeeld met het beoordelingsteam. De Boom LVS-toets Taalverzorging: Spelling is beoordeeld door een team bestaand uit één expert op het gebied van taal en één expert op het gebied van psychometrie. Daarnaast zijn twee editors van Stichting Cito betrokken bij de beoordeling, van wie één onderwijskundig editor en één psychometrisch editor. Op 5 november 2025 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 de hoor- en wederhoorgesprekken op 13 november 2025.

## 2 Algemene informatie

De algemene informatie over het leerlingvolgsysteem komt deels uit het document 'Boom LVS Taalverzorging: Spelling Handleiding'.

De Boom LVS-toets Taalverzorging: Spelling is onderdeel van het Boom leerlingvolgsysteem. De Boom LVS-toetsen zijn beschikbaar voor de kernvaardigheden rekenen-wiskunde, begrijpend lezen en taalverzorging, en voor de technische vaardigheden hoofdrekenen en technisch lezen.

De Boom LVS-toets Taalverzorging: Spelling is de opvolger van de Boom LVS-toets Spelling uit 2015. De toets meet de spellingvaardigheid van leerlingen zoals deze wordt aangeleerd in het Nederlandse basisonderwijs. Het LVS Taalverzorging bestaat uit lineaire papieren toetsen. Voor de heruitgave in 2025 was er volgens de aanbieder geen aanleiding om de toetsinhoud aan te passen. De aanbieder geeft aan dat de toetsen in volle tevredenheid door basisscholen ingezet worden. Daarnaast stelt de aanbieder dat zich geen cruciale ontwikkelingen in het spellingonderwijs hebben voorgedaan die reden tot aanpassing van de toetsen zou kunnen geven.

Naast de Boom LVS-toets Taalverzorging: Spelling werkt de uitgever aan een aanvullende Boom LVS-toets Taalverzorging: Grammatica en Interpunctie. Deze toetsen vormen samen één toetspakket om Taalverzorging te meten. Wanneer al deze toetsen zijn afgenomen bij leerlingen van groep 6 tot en met groep 8, kan het referentieniveau van deze leerlingen worden bepaald. Dit wordt besproken in de nog te verschijnen Gebruikershandleiding & Verantwoording van Boom LVS Taalverzorging: Grammatica & Interpunctie.

### Doelgroep en meetpretentie

De Boom LVS-toets Taalverzorging: Spelling Woorden is bestemd voor groep 3 tot en met groep 8 van het Nederlandse basisonderwijs. De Boom LVS-toets Taalverzorging: Spelling Werkwoorden is bestemd voor de bovenbouw van het basisonderwijs, dat wil zeggen, groep 6 tot en met groep 8. De Boom LVS-toets Taalverzorging: Spelling meet de spellingvaardigheid

van leerlingen zoals deze wordt aangeleerd in het Nederlandse basisonderwijs. Daarbij gaat het niet om een toetsing van de expliciete kennis van de spellingregels, maar om de toetsing van de expliciete dan wel impliciete beheersing van deze regels zoals die tot uitdrukking komt in het correct schrijven van woorden.

Het gebruiksdoel van de Boom LVS-toets Taalverzorging: Spelling is drieledig. Op de eerste plaats kan de toets op de basisschool worden ingezet voor niveaubepaling van de leerlingen, zowel individueel als per groep. Voor de niveaubepaling wordt elke behaalde vaardigheidsscore uitgedrukt in een normscore. Een tweede gebruiksdoel van de Boom LVS-toets Taalverzorging: Spelling is de voortgangsbepaling van de leerlingen. Wanneer de toets ten minste jaarlijks wordt afgenomen, kan deze worden ingezet ten behoeve van het leerlingvolgsysteem. Een derde gebruiksdoel van de Boom LVS-toets Taalverzorging: Spelling is het analyseren van fouten van leerlingen die moeite hebben met spellen.

### Inhoudelijke verantwoording

De Boom LVS-toets Taalverzorging: Spelling Woorden kan op een vast moment in het schooljaar afgenomen worden om de ontwikkeling in spellingvaardigheid vanaf groep 3 tot en met groep 8 te volgen; de toets Spelling werkwoorden kan worden afgenomen vanaf groep 6. Door verschillende meetmomenten te hanteren kunnen leerlingen met zichzelf of met de prestatie van een gemiddelde leerling met dezelfde onderwijservaring (de norm) vergeleken worden. Afwijkingen ten opzichte van een eerder vastgestelde ontwikkelingstrend of de norm kunnen hierdoor tijdig worden gesignaleerd. Om het niveau te bepalen wordt elke behaalde ruwe score uitgedrukt in een normscore. Het Boom LVS maakt hiervoor gebruik van verschillende normscores: niveau-indeling A t/m E en I t/m V, percentielscore en Didactisch Leeftijdsequivalent (DLE).

De toetsvorm is een dictee, waarbij leerlingen een zin te horen krijgen en ze vervolgens een woord uit de zin opschrijven. De leerlingen kunnen voor het correct schrijven van een woord gebruikmaken van de context van de zin waarin het woord wordt aangeboden. De Boom LVS-toets Taalverzorging: Spelling is opgebouwd uit een aansluitende reeks van toetsen ('blokken') van verschillend niveau. Naast het niveau wordt er een onderscheid gemaakt tussen het toetsen van woorden en het toetsen van werkwoorden, ook wel onveranderlijke respectievelijk veranderlijke woorden genoemd.

Het onderdeel Spelling Woorden bestaat uit zes blokken en het onderdeel Spelling Werkwoorden uit drie blokken. Een blok bestaat uit 15 tot 18 zinnen waarvan telkens een woord wordt opgeschreven. Er moeten per onderdeel altijd twee blokken per afname gemaakt worden door een leerling. De Boom LVS-toets Taalverzorging: Spelling kent per blok twee parallelle versies, die ontwikkeld zijn om de inzetbaarheid van het dictee in een doorlopend leerlingvolgsysteem te vergroten en bij herhaalde afnames bij een leerling een leereffect te voorkomen. De versies A en B zijn vergelijkbaar en kunnen afwisselend worden afgenomen. De leerkracht noteert de foutief geschreven woorden waarmee de score door de leerkracht vastgesteld kan worden met behulp van het Boom testcentrum.

De aanbieder heeft de volgende documentatie aangeleverd:

- Bijlage 1 - Leeswijzer Boom LVS Taalverzorging – Spelling woorden en werkwoorden
- Bijlage 2 - Boom LVS Taalverzorging Spelling Handleiding
- Bijlage 3 - Boom LVS ouderbrochure
- Bijlage 4 - Boom LVS Taalverzorging Spelling Dicteeboekje
- Bijlage 5 - Boom LVS Taalverzorging Spelling Toetsbladen
- Bijlage 6 - Boom LVS Map voor materialen Taalverzorging
- Bijlage 7 - Boom LVS Taalverzorging Spelling Instructiekaart voor de leerkracht
- Bijlage 8 - Gegevens proefaccount CvTE Boom testcentrum (digitale verwerking Boom LVS taalverzorging)
- Bijlage 9 - Boom LVS Stappenplan Boom testcentrum toets invoeren-rapport maken
- Bijlage 10 - ISO-certificaat
- Bijlage 11 - Toelichting bij het veilig bewaren van en omgaan met het materiaal voor scholen
- Bijlage 12 - Leerlingrapport LVS Taalverzorging Spelling
- Bijlage 13 - Groepsrapport LVS Taalverzorging
- Bijlage 14 - Excelbestand Boom LVS Taalverzorging Spelling Werkwoorden normtabel
- Bijlage 15 - Excelbestand Boom LVS Taalverzorging Spelling Woorden normtabel

### 3 Kwaliteitscriteria beoordelingskader

De Boom LVS-toets Taalverzorging: Spelling is een bestaand 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	Hoofdstuk	Kwaliteitseis
Onderwijskundige aspecten	2	O.III t/m O.VI
Psychometrische aspecten	3	P.IV t/m P.VIII
Organisatorische aspecten	4	B.I

Hieronder volgt een kort overzicht van de belangrijkste bevindingen van het beoordelingsteam op de onderwijskundige, organisatorische en psychometrische aspecten. Een overzicht van de beoordeling per kwaliteitseis is te vinden in Bijlage 1.

#### Onderwijskundige aspecten

De Kwaliteitseisen O.III.2, O.IV.1, O.IV.2, O.V.1b waren door het beoordelingsteam in het conceptadvies beoordeeld als onvoldoende onder voorbehoud. Alle andere eisen waren als voldoende beoordeeld.

De aanbieder toont in de verantwoording van de kwaliteitscriteria behorende bij paragraaf 2.3 aan dat de toets voldoet aan de gestelde eisen. De uitgangspunten van de toetsconstructie

zijn als voldoende beoordeeld (O.III.1 tot en met O.III.5). De Boom LVS-toets Taalverzorging: Spelling voor groep 3 tot en met 8 meet de spellingvaardigheid aan de hand van dictees. De opbouw en de inhoud van de dicteeblokken zijn geschikt en relevant voor de doelgroep. De A- en B-versie van de dictees zijn gelijk qua opbouw van gemakkelijk naar moeilijk en bieden de leerkrachten de mogelijkheid om de toets in te zetten voor niveaubepaling, voortgangsbepaling en het signaleren van spellingsmoeilijkheden. De afnamevorm, de doelgroep, het gebruiks- en meetdoel en de theoretische inkadering zijn uitvoerig en adequaat geformuleerd. Met betrekking tot de doelgroep (O.III.2) heeft het beoordelingsteam de aanbieder verzocht informatie aan te leveren in welke situatie(s) en/of bij welke individuele ondersteuningsbehoefte het instrument niet geschikt is.

De verantwoording van de kwaliteitseisen rondom de inhoudsvaliditeit is voldoende bevonden (O.IV tot en met O.IV.3). De toetsmatrijs is een adequate representatie van het meetdoel (O.IV.1). De opgaven in de vorm van dicteewoorden die in zinnen aangeboden worden, voldoen aan de constructievoorschriften en eisen uit de checklist voor het beoordelen van de kwaliteit van observatie-categorieën en toetsopgaven. Het beoordelingsteam is van mening dat de gebruikte woorden passen bij de doelgroep van het instrument (O.IV.2). Bij een aantal doelwoorden in de toetsblokken voor de onderbouw heeft het beoordelingsteam de aanbieder gevraagd te reflecteren op opmerkingen over de contextzinnen en waar nodig de opgaven aan te passen. Het beoordelingsteam miste een uitgebreidere verslaglegging van de actualiteitscontrole van de toets. Daarnaast heeft het beoordelingsteam tijdens het hoor- en wederhoorgesprek aan de aanbieder gevraagd te reflecteren op de mogelijke relatie tussen het risico op itemexposure en de uit de psychometrische analyses van de toetsen gebleken significant toegenomen spellingvaardigheid (O.IV.2).

Het beoordelingsteam heeft bij kwaliteitseis O.V.1b geconstateerd dat er bij Spelling Woorden voor de bovenbouw in verhouding meer 1F opgaven dan 2F (en 3F) opgaven zijn opgenomen. Het beoordelingsteam heeft de aanbieder gevraagd om hierop te reflecteren.

De kwaliteitseisen die betrekking hebben op inzicht in de leervorderingen zijn als voldoende beoordeeld (O.VI.1 tot en met O.VI.4). De leerlingrapporten zijn duidelijk vormgegeven en de informatie is helder verwoord. De ouderbrochure biedt voor ouders, verzorgers, voogden en docenten een toelichting die begrijpelijk is.

De genoemde punten bij kwaliteitseisen O.III.2, O.IV.1, O.IV.2 en O.V.1b zijn ter sprake gebracht tijdens het hoor- en wederhoorgesprek.

## Psychometrische aspecten

De aanbieder heeft informatie gepresenteerd over de steekproef. Deze voldoet aan de vereiste aantallen. De aanbieder heeft de representativiteit in kaart gebracht. In eerste instantie waren niet alle tabellen in orde, wat de interpretatie van de gepresenteerde resultaten bemoeilijkte en de aanbieder werd gevraagd correcte statistieken op te leveren (P.IV.1). De aanbieder heeft het afnamesdesign helder uiteengezet. Het beoordelingsteam had in eerste instantie wat twijfels over de gebruikte kalibratiemethode en de aanbieder werd gevraagd om opheldering. De itemparameters verschillen in de regel niet tussen de

toetsversies, indien er wel verschillen waren, zijn de parameters van deze items los geschat. Er is ook onderzoek naar DIF gedaan ten aanzien van schoolweging. Er waren enkele items die DIF lieten zien en deze items worden gehandhaafd, wat het beoordelingsteam gerechtvaardigd vindt, wel verzocht het beoordelingsteam nog om effect sizes. De modelpassing is zowel gevisualiseerd met ICC's als aan de hand van een effect size. Het beoordelingsteam onderschreef dat het model voldoende goed bij de data past. De meetnauwkeurigheid is per toetsboekje voldoende voor de verschillende vaardigheidsscores. De aanbieder werd nog gevraagd hoe wordt omgegaan met vermoeidheids- en/of volgorde-effecten (P.IV.2). Er is sprake van een continue relatieve normering. De normeringssteekproef is voldoende groot. (P.V.4). De aanbieder evalueerde de representativiteit van de normeringssteekproef in eerste instantie aan de hand van een minder geschikte methode en ze werd gevraagd om een passende methode te gebruiken (P.V.5). De normen lijken correct te zijn bepaald (P.V.6). De aanbieder heeft beschreven hoe ze de globale betrouwbaarheid op de vaardigheidsschaal berekent. Dit werd niet beschreven voor de meetprecisie en de aanbieder werd verzocht dit alsnog te doen (P.VI.1). De betrouwbaarheidsindices zijn bevredigend (P.VI.2). De aanbieder werd verzocht opheldering te verschaffen rondom de analyses van de percentages correcte (mis)classificaties (P.VI.3). De unidimensionaliteit is voldoende aangetoond (P.VII.1). De aanbieder werd gevraagd de klassieke itemstatistieken in meer detail te verantwoorden. Ten behoeve van de dekking van de toetsmatrijs heeft de aanbieder in sommige gevallen besloten om items in de toets te laten (P.VII.2). De convergente en divergente validiteit is volgens het beoordelingsteam voldoende aangetoond (P.VII.3). Er werden enkele items met DIF gevonden ten aanzien van stedelijkheid en deze items zijn gehandhaafd, wat het beoordelingsteam gerechtvaardigd vindt. Ook ten aanzien van dit punt werd de aanbieder verzocht om effect sizes aan te leveren (P.VII.4). De prestaties van relevante groepen zijn vergeleken. De aanbieder werd verzocht om extra duiding te geven aan de gevonden verschillen (P.VII.5). De groei wordt op een adequate manier gemeten (P.VIII.1). Er werden in eerste instantie geen indicaties gegeven van meetprecisie van verschilcores en de aanbieder is gevraagd dit alsnog te doen (P.VIII.2). Er zijn voldoende en duidelijke gegevens verstrekt hoe de groei geïnterpreteerd dient te worden (P.VIII.3).

De genoemde punten zijn ter sprake gebracht tijdens het hoor- en wederhoorgesprek.

### Organisatorische aspecten

De aanbieder is ISO 27001 gecertificeerd met een geldigheid tot 12-05-2026. De kwaliteitscriteria in paragraaf 4.1 zijn door de ISO-certificering voldoende verantwoord.

## 4 Hoor- en wederhoorgesprek

Op 25 en 27 november 2025 hebben onder regie van het CvTE de hoor- en wederhoorgesprekken plaatsgevonden. Vanuit het beoordelingsteam zijn kwaliteitseisen O.III.2, O.IV.1, O.IV.2, O.V.1b, P.IV.1, P.IV.2, P.V.4, P.V.5, P.VI.1, P.VI.3, P.VII.2, P.VII.4, P.VII.5 en P.VIII.2 ter sprake gebracht, omdat het beoordelingsteam informatie miste om deze eisen te kunnen beoordelen. Ook de andere aandachtspunten in Bijlage 1 zijn ter verduidelijking van de aangeleverde informatie besproken.

## 5 Proces richting definitief advies

Op 10 december 2025 is de gevraagde informatie in een nazending door de aanbieder aangeleverd. De editors van Stichting Cito hebben de aanvullende informatie beoordeeld en hebben vastgesteld dat de verantwoording op alle kwaliteitseisen voldoende is.

Het beoordelingsteam heeft de aanbieder gevraagd informatie aan te leveren in welke situatie(s) en/of bij welke individuele ondersteuningsbehoefte het instrument niet geschikt is (O.III.2). De aanbieder heeft in de nazending uitgebreid en adequaat vermeld welke mogelijkheden wel of niet geboden worden voor leerlingen met een aantal specifieke ondersteuningsbehoefte zoals slechtziende leerlingen, blinde leerlingen, leerlingen met een motorische beperking of leerlingen met een auditieve beperking.

De aanbieder heeft op verzoek van het beoordelingsteam alle relevante tabellen voor het verantwoorden van kwaliteitseis O.IV.1 gegroepeerd, zodat de beoordelaar sneller de vereiste informatie (doelwoord, categorie volgens het Referentiekader/SLO-indeling, referentieniveau) kan afleiden. Op basis van deze tabellen concludeert het beoordelingsteam dat de toetsmatrijs een adequate, schematische representatie van het meetdoel is.

Bij een aantal doelwoorden in de toetsblokken voor de onderbouw heeft het beoordelingsteam de aanbieder gevraagd te reflecteren op opmerkingen over de contextzinnen en waar nodig de opgaven aan te passen (O.IV.2). In de aanvullende informatie toont de aanbieder dat de contextzinnen bij vier opgaven zijn aangepast, waardoor de betekenis van het doelwoord duidelijker of minder gevoelig is. Daarnaast is een doelwoord vervangen door een gangbaarder woord met dezelfde spellingsregel. Op basis van de doorgevoerde wijzigingen concludeert het beoordelingsteam dat alle toetsopgaven voldoen aan de gestelde kwaliteitseisen in de Checklist voor het beoordelen van de kwaliteit van observatiecategorieën en toetsopgaven.

Daarnaast heeft het beoordelingsteam aan de aanbieder gevraagd om een adequate verantwoording te geven van de actualiteitscontrole. In de aanvullende informatie een uitgebreidere en navolgbare toelichting met betrekking tot de actualiteitscheck gepresenteerd. Tot slot heeft de aanbieder met betrekking tot O.IV.2 toegelicht waarom itemexposure een aanvaardbaar risico is. Op basis hiervan concludeert het beoordelingsteam dat de huidige toetsamenstelling voldoet aan de kwaliteitseis.

Met betrekking tot kwaliteitseis O.V.1b heeft de aanbieder een toelichting gegeven op de verhouding 1F, 2F, 3F opgaven in de toets Spelling Woorden voor de bovenbouw. Op basis van deze argumentatie concludeert het beoordelingsteam dat er voldoende opgaven op de niveaus 2F en 3F in de toetsen zijn opgenomen om een onderbouwde uitspraak te doen over het behaalde niveau.

De aanbieder heeft in de nazending alle gevraagde informatie ten aanzien van de onderwijskundige aspecten aangeleverd en alle eisen als voldoende kunnen beoordelen.

Met betrekking tot de psychometrische eisen heeft de aanbieder ook aanvullende informatie aangeleverd.

De aanbieder heeft in eerste instantie Cramer's V aangeleverd met incorrecte interpretatiecriteria. In de aanvullende informatie is Cohen's W berekend en zijn de resultaten geëvalueerd met de correcte cut-off-scores. De steekproef was voldoende representatief (P.IV.1). De aanbieder voert de kalibratie uit met de JML-schatter, waarbij niet de itemparameters van alle toetsversies in één concurrente kalibratie zijn geschat. Dit leek het beoordelingsteam inefficiënt en JML wordt tegenwoordig nog weinig gebruikt. De aanbieder liet zien dat het in de praktijk niet uitmaakt of hun procedure wordt gehanteerd of een concurrente kalibratie met CML voor deze toepassing. Ze heeft effect sizes aangeleverd bij de DIF-analyses en ook heeft de aanbieder geëxpliciteerd hoe om wordt gegaan met eventuele vermoeidheids- en/of volgorde-effecten (P.IV.2). De aanbieder heeft – net als bij P.IV.1 - een passende maat voor representativiteitsanalyse gekozen (P.V.5). De berekening van de meetprecisie is verduidelijkt (P.VI.1) en de indices voor de misclassificaties zijn toegelicht (P.VI.3), en voldoende geacht door het beoordelingsteam. De klassieke itemindices zijn op itemniveau aangeleverd en de aanbieder heeft duidelijk beargumenteerd waarom soms wordt afgeweken van de eigen opgestelde cut-off criteria (P.VII.2). DIF-analyses zijn opnieuw uitgevoerd waarbij ook effect sizes zijn aangeleverd (P.VII.4). De gevonden groepsverschillen zijn geduid (P.VII.5). Indicaties van meetprecisie van verschilscores zijn gegeven (P.VIII.2).

Met de aanvullende informatie heeft het beoordelingsteam alle eisen als voldoende kunnen beoordelen.

## 6 Oordeel en advies

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

Paragraaf	Kwaliteitseis	Codering kwaliteitseis	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

Paragraaf	Kwaliteitseis	Codering kwaliteitseis	Oordeel
4.1	Afname	B.I	Voldoende

Stichting Cito geeft op basis van de verstrekte verantwoordingsdocumenten een positief advies voor de tussentijdse check van de Boom LVS-toets Taalverzorging: Spelling Woorden en Spelling Werkwoorden die onderdeel uitmaakt van het Boom leerlingvolgsysteem.

Dit advies is vastgesteld op 18 december 2025 te Arnhem.

## Bijlage 1: Toelichting beoordeling per kwaliteitscriterium

### 2 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?	<p>De afnamevorm van het Boom LVS Taalverzorging: Spelling is gedegen en compleet gedefinieerd. De toets is een dictee op papier en is zowel per groep als individueel af te nemen. Daarbij gaat het om het correct schrijven van losse woorden.</p> <p>De leerkracht leest de dicteezinnen voor en geeft aan welk woord er moet worden opgeschreven of ingevuld op het antwoordblad (zie de documenten 'Bijlage 4 - Boom LVS Taalverzorging Spelling Dicteeboekje' en 'Bijlage 5 - Boom LVS Taalverzorging Spelling Toetsbladen').</p> <p>De Boom LVS-toets Taalverzorging: Spelling Woorden en Spelling Werkwoorden bestaat uit meerdere blokken. Het onderdeel Spelling Woorden bestaat uit zes blokken van 15 opgaven. Spelling Werkwoorden bestaat uit drie blokken van 18 opgaven. Er moeten altijd per onderdeel twee blokken worden gemaakt. Bij de Boom LVS-toets Taalverzorging: Spelling Woorden bestaat elke toets uit 30 opgaven, bij de Boom LVS-toets Taalverzorging: Spelling Werkwoorden bestaat elke toets uit 36 opgaven. Voor het maken van de opgaven van twee blokken volstaat een tijd van ongeveer 30 minuten.</p>	Ja

		Bij de indeling van de opgaven in blokken is zo goed mogelijk aangesloten bij het moment waarop bepaalde spellingregels in het onderwijsaanbod aan de orde komen. Zie het document 'Bijlage 2 - Boom LVS Taalverzorging: Spelling Handleiding', Hoofdstuk 3: Inleiding, paragraaf 3.1 en paragraaf 3.3.	
O.III.2	Is (zijn) de doelgroep(en) van het instrument gedegen en compleet gedefinieerd? Of is er expliciet benoemd voor welke doelgroep(en) het instrument niet geschikt is?	<p>De doelgroepen van het instrument zijn gedegen en compleet gedefinieerd. De Boom LVS-toets Taalverzorging: Spelling Woorden is bestemd voor groep 3 tot en met groep 8 van het Nederlandse basisonderwijs; de Boom LVS-toets Taalverzorging: Spelling Werkwoorden is bestemd voor de bovenbouw van het basisonderwijs, dat wil zeggen, groep 6 tot en met groep 8. Zie 'Bijlage 2 - Boom LVS Taalverzorging: Spelling Handleiding', Hoofdstuk 3: paragraaf 3.1 en Hoofdstuk 1, paragraaf 1.2 en 1.6. In hetzelfde document wordt in paragraaf 3.2 beschreven voor welke groepen de verschillende blokken bedoeld zijn.</p> <p>Het dictee wordt klassikaal afgenomen. In sommige gevallen kan er gekozen worden voor een individuele afname. In Bijlage 2 (paragraaf 1.6, pagina 19) staat aangegeven hoe de toets individueel af te nemen is bij leerlingen die moeite hebben met spellen.</p> <p>De LVS-toets is geschikt voor leerlingen met (een vermoeden van) ernstige enkelvoudige dyslexie (EED). In het document 'Bijlage 2 - Boom LVS Taalverzorging Spelling Handleiding' adviseert de aanbieder om voor deze doelgroep bij een individuele afname andere (lagere) blokken af te nemen om frustratie te voorkomen en de informatieve waarden van de resultaten te vergroten. Er zijn verschillende mogelijkheden om tot een globale inschatting te komen van het spellingniveau van leerlingen en daarmee te bepalen welke blokken</p>	Ja

		<p>het meest geschikt zijn om af te nemen. Dit kan bijvoorbeeld met behulp van een eerdere toetsafname.</p> <p>Voor slechtziende leerlingen is vergroot materiaal beschikbaar voor zowel Spelling Woorden als Spelling Werkwoorden. Omdat de leerling bij het onderdeel Spelling Werkwoorden het werkwoord in een (voorgedrukte) zin moet invullen (de context is daarbij immers van belang), adviseert de aanbieder het vergrote materiaal te gebruiken.</p> <p>Bij de reguliere versie schrijven leerlingen de antwoorden op een invulblad. Voor leerlingen met een motorische beperking is het mogelijk de woorden te typen en in te leveren bij de leerkracht. Hierbij is het zaak dat de spellingscontrole op de computer uitgeschakeld is.</p> <p>Op dit moment is er voor blinde leerlingen geen gebrailleerd materiaal beschikbaar. Daarbij zijn de toetsen niet vertaald naar Nederlandse Gebaren Taal (NGT) of Nederlands met Gebaren (NmG). Voor dove of blinde leerlingen is de toets dus niet geschikt.</p> <p>Het beoordelingsteam beoordeelt deze kwaliteitseis als voldoende.</p>	
O.III.3	Is het meetdoel van het instrument gedegen en compleet gedefinieerd?	<p>Het meetdoel van het instrument is gedegen en compleet gedefinieerd. De Boom LVS-toets Taalverzorging: Spelling meet de spellingvaardigheid van leerlingen zoals deze wordt aangeleerd in het Nederlandse basisonderwijs. Daarbij gaat het niet om het toetsen van de expliciete kennis van de spellingregels, maar om de toetsing van de expliciete dan wel impliciete beheersing van deze regels zoals die tot uitdrukking komt in het correct schrijven van woorden. De toetsvorm is een dictee, waarbij leerlingen een zin te horen krijgen en vervolgens aan hen wordt gevraagd een woord uit de zin op te schrijven. De leerlingen kunnen dus voor het correct schrijven van een woord</p>	Ja

		<p>gebruikmaken van de context van de zin waarin het woord wordt aangeboden.</p> <p>De LVS-toets bestaat uit verschillende blokken. Bij het samenstellen van de blokken is rekening gehouden met het moment waarop een spellingprobleem aan bod komt binnen het basisonderwijs. Zie het document 'Bijlage 2 - Boom LVS Taalverzorging: Spelling Handleiding', Hoofdstuk 1, paragraaf 1.3 en in hetzelfde document 'Bijlage II'. In 'Bijlage II' staan meerdere tabellen, waarbij de opgaven worden gekoppeld aan de categorieën (zoals weergegeven in Bijlage I).</p>	
O.III.4	Is er een actuele gedegen en gedetailleerde inhoudelijke theoretische inkadering van het instrument?	<p>Het instrument is voorzien van een actuele gedegen en gedetailleerde inhoudelijke theoretische inkadering, zie 'Bijlage 2 - Boom LVS Taalverzorging: Spelling Handleiding', Hoofdstuk 1 (paragraaf 1.4.3 t/m paragraaf 1.4.5) en Hoofdstuk 2 (paragraaf 2.2).</p> <p>De volgende zaken worden beschreven:</p> <ul style="list-style-type: none"> <li>• de verschillende fasen in het leren van het spellen (aanvankelijk en gevorderd);</li> <li>• welke instructiemethoden er zijn voor het aanleren van de spellingregels en -strategieën;</li> <li>• welke spellingproblemen voorkomen bij leerlingen;</li> <li>• verschillende onderzoeken naar spellingvaardigheid met betrekking tot signaleren en diagnosticeren;</li> <li>• de relatie van de spellingtoetsen met het Referentiekader taal en rekenen (2009) en CED-leerlijnen spelling.</li> </ul>	Ja
O.III.5	Is het gebruiksdoel van het instrument gedegen en compleet gedefinieerd?	<p>Het gebruiksdoel is gedegen en compleet gedefinieerd. Het gebruiksdoel van de Boom LVS-toets Taalverzorging: Spelling is drieledig. Op de eerste plaats kan de toets op de basisschool worden</p>	Ja

		<p>ingezet voor de niveaubepaling van leerlingen, zowel individueel als per groep. Voor de niveaubepaling wordt elke behaalde vaardigheidsscore uitgedrukt in een normscore. Het tweede gebruiksdoel van de Boom LVS-toets Taalverzorging: Spelling is de voortgangsbepaling van leerlingen. Wanneer de toets ten minste jaarlijks wordt afgenomen, kan deze worden ingezet ten behoeve van het leerlingvolgsysteem. Het derde gebruiksdoel van de Boom LVS-toets Taalverzorging: Spelling is het analyseren van het soort fouten die leerlingen gemaakt hebben tijdens het spellen. Zie het document 'Bijlage 1 - Boom LVS Taalverzorging: Spelling Handleiding', Hoofdstuk 3, paragraaf 3.4 en paragraaf 3.8.</p>	
--	--	---	--

### 2.3.2 Inhoudsvaliditeit

Code	Vragen	Toelichting	Antwoord
O.IV.1	Is de toetsmatrijs een adequate, schematische representatie van het meetdoel?	<p>Uit de aangeboden informatie is af te leiden dat de toetsmatrijs een adequate schematische representatie van het meetdoel is. In het document 'Bijlage 2 - Boom LVS Taalverzorging: Spelling Handleiding', Hoofdstuk 2, paragraaf 2.2 wordt beschreven op welke wijze de toets is opgebouwd.</p> <p>De toets bestaat per onderdeel uit twee blokken. In totaal 30 items Spelling Woorden en 36 items Spelling Werkwoorden. Er is één punt per vraag te behalen met een totaalscore van 30 op Spelling Woorden en 36 punten op Spelling Werkwoorden. De toetsduur is twee keer 30 minuten (Zie paragraaf 3.1, pagina 32).</p> <p>In 'Bijlage II' (pagina 100) van de Handleiding geven Tabel 1 en Tabel 2 per toetsblok een indeling op hoofdcategorie weer. Daarnaast wordt</p>	Ja

		<p>in Tabel 3 tot en met 6 per opgave aangegeven welke spellingscategorieën de opgave bevat. Op pagina 23 tot en met 26 wordt in Tabel 2.1 een overzicht gegeven van alle SLO-doelen en hoe frequent deze voorkomen in de toetsen van de verschillende leerjaren.</p> <p>In de aanvullende informatie heeft de aanbieder op verzoek van het beoordelingsteam alle relevante tabellen gegroepeerd, zodat de beoordelaar sneller de vereiste informatie (doelwoord, categorie volgens het Referentiekader/SLO-indeling, referentieniveau) kan afleiden. Het beoordelingsteam constateert dat de aanbieder bij het samenstellen van de dictees gangbare woorden heeft geselecteerd die een afspiegeling zijn van de diverse spellingsproblemen die zich in de Nederlandse taal voordoen en daarmee aansluiten op de passieve en actieve woordenschat van leerlingen in het basisonderwijs. Op basis van deze informatie concludeert het beoordelingsteam dat de toetsmatrijs een adequate, schematische representatie van het meetdoel is.</p>	
O.IV.2	Voldoen alle toetsopgaven in het instrument aan de kwaliteitscriteria voor toetsopgaven: relevantie, objectiviteit, efficiëntie, specificiteit en neutraliteit?	<p>De toetsopgaven voldoen aan de kwaliteitscriteria uit Checklist 1: relevantie, objectiviteit, efficiëntie, specificiteit en neutraliteit.</p> <p>Het gaat om zinnen voor een dictee. In het kader van dat doel is het laten opschrijven van een woord uit een losse zin relevant. De dicteezinnen voor de A en B versies van het dictee, zoals weergegeven in het document 'Bijlage 4 - Boom LVS Taalverzorging Spelling Dicteeboekje', zijn gecontroleerd aan de hand van bovengenoemde kwaliteitscriteria. De getoetste woorden passen bij de bedoelde categorieën en contextzinnen zijn adequaat.</p> <p>Bij een aantal doelwoorden en contextzinnen in de toetsblokken voor de onderbouw vroeg het beoordelingsteam zich af of de gebruikte</p>	Ja

		<p>woorden passend zijn bij de beoogde doelgroep. De aanbieder heeft de contextzinnen bij vier opgaven aangepast, waardoor de betekenis van het doelwoord duidelijker of minder gevoelig is. Daarnaast is een doelwoord vervangen door een gangbaarder woord met dezelfde spellingsregel.</p> <p>De toets bevat geen nieuwe opgaven en is sinds 2015 niet meer aangepast. De aanbieder geeft aan dat er geen aanleiding was om de toetsen aan te passen: 'De toetsen worden in volle tevredenheid door basisscholen ingezet en hebben zich in de tussentijd bewezen. Er hebben zich ook geen cruciale ontwikkelingen in het spellingonderwijs voorgedaan die een reden zouden kunnen zijn om de toetsen aan te passen.'</p> <p>In de aanvullende informatie geeft de aanbieder op verzoek van het beoordelingsteam een uitgebreidere toelichting met betrekking tot de actualiteitscheck. De aanbieder stelt dat actualiteit van woorden in een spellingtest vooral relevant is wanneer je de spelling van woorden uit het hoofd gaat leren (intensieve herhaling, overleren). De aanbieder gaat echter uit van de regelgerichte leermethodiek met expliciete systematische instructie. Hierbij oefenen leerlingen geen losse (actuele) woorden, maar worden spellingprincipes aangeleerd waarmee je actuele én minder-actuele woorden kunt spellen. Een uitgebreide actualiteitscheck op woordniveau acht de aanbieder hierdoor niet nodig. Het beoordelingsteam constateert dat de uitgebreidere en navolgbare toelichting met betrekking tot de actualiteitscheck adequaat en voldoende is. Het beoordelingsteam heeft op basis van deze argumentatie geconcludeerd dat de huidige contextzinnen en doelwoorden passend zijn.</p> <p>Het beoordelingsteam heeft de aanbieder tijdens het hoor- en wederhoorgesprek gevraagd te reflecteren op de mogelijke relatie</p>	
--	--	---	--

		<p>tussen het risico op itemexposure en de uit de psychometrische analyses van de toetsen gebleken significant toegenomen spellingvaardigheid. Doordat de toetsen al meerdere jaren op scholen liggen, kan er sprake zijn van een risico op exposure.</p> <p>Volgens de aanbieder is exposure niet echt een risico. In de regel geldt, hoe meer exposure van spelling des te beter leerlingen de spelling leren. Exposure uit een spellingtest die een jaar eerder is gemaakt, is zeer onwaarschijnlijk. Naar verwachting zijn de leerlingen in de tussentijd nog meerdere malen blootgesteld aan de spelling van dat woord. De aanbieder vergelijkt het met rekensommen: ook hier worden leerlingen vaker aan dezelfde sommen blootgesteld en komen deze sommen ook voor in toetsen om de rekenvaardigheid te meten.</p> <p>De reden dat de spellingvaardigheid is toegenomen kan de aanbieder niet direct verklaren. Duidelijk is wel dat zogenoemde exposure hier met aangrenzende waarschijnlijkheid niet de oorzaak van is. De aanbieder ziet ook dat de spellingvaardigheid is toegenomen, maar dat de items geen eenduidig beeld vormen. Niet bij ieder item is de vaardigheid van de leerlingen toegenomen.</p> <p>Op basis hiervan concludeert het beoordelingsteam dat de huidige toetsamenstelling voldoet aan de gestelde eis in het Beoordelingskader.</p>	
O.IV.3	Voldoen de toetsopgaven aan de constructievoorschriften passend bij het vraagtype?	<p>De toetsopgaven in het instrument voldoen aan de constructievoorschriften zoals gedefinieerd in Checklist 1.</p> <p>Zowel bij het onderdeel Spelling Woorden als bij het onderdeel Spelling Werkwoorden is sprake van een dictee waarbij eerst een zin wordt voorgelezen. Daarna wordt het woord herhaald dat opgeschreven moet worden. Zo is de context van het woord duidelijk.</p>	Ja

		Het antwoordmodel is eenduidig. De dicteezinnen voor de A en B versies van het dictee zoals weergegeven in het document 'Bijlage 4 - Boom LVS Spelling Dicteeboekje' zijn gecontroleerd aan de hand van kwaliteitscriteria in de Checklist. De getoetste woorden passen bij de bedoelde categorieën. De uitgangspunten voor de toetsconstructie zijn uitgebreid beschreven. Zie 'Bijlage 2 - Boom LVS Taalverzorging: Spelling Handleiding', Hoofdstuk 2, paragraaf 2.2.	
--	--	--	--

### 2.3.3 Referentieniveaus

Code	Vragen	Toelichting	Antwoord
O.V.1a	Worden de wettelijk verplichte inhouden van de terreinen Nederlandse taal en/of Rekenen, en indien van toepassing de optionele inhouden van Taal, evenwichtig getoetst?	De wettelijk verplichte inhouden van het terrein Nederlandse taal bij het subdomein Spelling (niet-werkwoorden en werkwoorden) worden evenwichtig getoetst. In 'Bijlage 2 - Boom LVS Taalverzorging Spelling Handleiding', worden in Tabel 2.1 de spellingcategorieën naar SLO-Meijerink weergegeven. Hier is te zien dat deze categorieën aan bod komen in de verschillende spellingblokken en evenwichtig verdeeld zijn. In paragraaf 2.4 (pagina 29) is de verdeling van de opgaven over de referentieniveaus inzichtelijk gemaakt.	Ja

O.V.1b	<p>Zijn de opgaven van de toets voor de wettelijk verplichte inhouden van Nederlandse taal en/of Rekenen, en indien van toepassing de optionele inhouden van Taal, evenwichtig samengesteld qua referentieniveaus 1F en/of 2F/1S?</p>	<p>Het LVS-instrument bevat voor wat betreft het wettelijk verplichte voorgeschreven onderdeel Taalverzorging van het terrein Nederlandse taal bij het subdomein Spelling (niet-werkwoorden en werkwoorden) de inhouden van 1F, 2F en 3F. Er heeft bij de dicteewoorden een inhoudelijk correcte toekenning van referentieniveaus plaatsgevonden (zie paragraaf 1.4.4 en 2.4 in 'Bijlage 2 - Boom LVS Taalverzorging: Spelling Handleiding'.</p> <p>De toetsen Woorden en Werkwoorden zijn voldoende evenwichtig samengesteld qua referentieniveaus 1F, 2F en/of 3F. In verhouding zijn er meer opgaven 1F dan 2F (en 3F) opgenomen, met name bij Spelling Woorden. Het beoordelingsteam heeft de aanbieder gevraagd om hierop te reflecteren. De aanbieder geeft in de aanvullende informatie aan dat de theoretische indeling van de SLO-doelen geen goede weergave zijn van de complexiteit van de getoetste spellingstof. De opgaven uit deze blokken die een 1F categorisering hebben gekregen, zullen vaak pas goed worden geschreven door leerlingen die minimaal een 2F spellingniveau hebben. Het beoordelingsteam kan deze argumentatie volgen en concludeert dat er voldoende opgaven op de niveaus 2F en 3F items in de toetsen zijn opgenomen om een onderbouwde uitspraak te doen over het behaalde niveau.</p> <p>Hierbij wil het beoordelingsteam nog aangeven dat in het te verschijnen deel 2 van het Boom LVS Taalverzorging: Grammatica en Interpunctie de koppeling met de referentieniveaus gemaakt zal worden. Voor deze bepaling zal een cesuur voor het Boom LVS Taalverzorging: Spelling Woorden, Spelling Werkwoorden, Grammatica en Interpunctie worden vastgesteld. Daarbij zal gebruikgemaakt worden van de openbare referentieset/ankerset Taalverzorging 1F en 2F, die Stichting Cito – onder</p>	Ja
--------	---	---	----

Code	Vragen	Toelichting	Antwoord
		verantwoordelijkheid van het CvTE en in opdracht van het ministerie van OCW – heeft ontwikkeld. Door de toets met deze ankersets te verankeren zal een cesuur bepaald worden waarmee kan worden vastgesteld of een leerling het fundamentele niveau 1F dan wel het fundamentele niveau of het streefniveau 2F heeft gehaald op het subdomein Taalverzorging.	
O.V.2a	Indien er sprake is van een CAT (adaptieve toets op itemniveau), voldoet de itembank van de verzameling van LVS-instrumenten van type (III) van (eind) groep 6 tot en met 8 aan de kwaliteitseis O.V.1a en O.V.1b?	Er is geen sprake van een CAT dus deze kwaliteitseis is niet van toepassing.	N.v.t.
O.V.2b	Indien er sprake is van een MST (adaptieve toets op moduleniveau), voldoet het design van de verzameling van LVS-instrumenten van het type (III) van (eind) groep 6 tot en met 8 aan de kwaliteitseis O.V.1a en O.V.1b?	Er is geen sprake van een MST dus deze kwaliteitseis is niet van toepassing.	N.v.t.

### 2.3.4 Inzicht in leervorderingen

Code	Vragen	Toelichting	Antwoord
O.VI.1	Stelt de toetsaanbieder voor elke individuele leerling een leerlingrapport op?	De aanbieder stelt voor elke leerling een leerlingrapport op. Het leerlingrapport bevat de score, het percentiel, niveau I-V, niveau A-E en een DLE (zie 'Bijlage 12 - Leerlingrapport LVS Taalverzorging Spelling').  In het document 'Bijlage 2 - Boom LVS Taalverzorging Spelling' geeft de aanbieder in Hoofdstuk 3.5 een toelichting bij het leerlingrapport.	Ja
O.VI.2	Bevat het leerlingrapport het gemeten niveau van de leerling en geeft het rapport de leervordering van de leerling weer?	In 'Bijlage 12 - Leerlingrapport LVS Taalverzorging Spelling' is te zien dat op het leerlingrapport het gemeten niveau wordt gerapporteerd en de leervordering wordt weergegeven. Er wordt bijvoorbeeld een vaardigheidsscore gerapporteerd, zie ook O.VI.1.  In 'Bijlage 2 - Boom LVS Taalverzorging Spelling' wordt in Hoofdstuk 3.6 beschreven dat er na de afname van de LVS-toets per leerling een Individueel overzicht opgevraagd kan worden in het Boom Testcentrum. Dit overzicht bevat een historisch overzicht van alle toetsresultaten. Door deze resultaten naast elkaar te zetten, kan de groei van een leerling in beeld worden gebracht.	Ja
O.VI.3	Biedt de aanbieder een gedegen en begrijpelijke toelichting bij het gemeten niveau en de leervorderingen van de leerling, zodat leerkrachten deze adequaat kunnen interpreteren en toepassen?	De aanbieder biedt een gedegen en begrijpelijke toelichting bij het gemeten niveau en de leervorderingen van de leerling waarmee de leerkracht de resultaten adequaat kan interpreteren en toepassen. Zie het document 'Bijlage 2 - Boom LVS Taalverzorging: Spelling Handleiding', Hoofdstuk 3, paragraaf 3.4 t/m 3.6.	Ja

Code	Vragen	Toelichting	Antwoord
O.VI.4	Stelt de aanbieder begrijpelijk en toegankelijk informatie aan leerkrachten beschikbaar over het gemeten niveau en de leervorderingen van de leerling, zodat zij ouders, verzorgers, en voogden voldoende en volledig kunnen informeren?	Het document 'Bijlage 3 - Boom LVS ouderbrochure' bevat voor ouders en verzorgers een begrijpelijke toelichting over het gemeten niveau en de leervorderingen van de leerling. In de brochure staat algemene informatie over het doel en de toetsvorm van de LVS-toetsen. Daarnaast wordt beschreven hoe de resultaten geïnterpreteerd kunnen worden. Moeilijke begrippen, zoals DLE, vaardigheidsscore en referentieniveau, worden uitgelegd in begrijpelijke taal. De brochure bevat ook een voorbeeldrapportage.	Ja
O.VI.5	Is de informatie in het leerlingrapport een voldoende basis voor de leerkracht om de leervorderingen te evalueren en vervolgstappen te formuleren?	De informatie in het leerlingrapport is een voldoende basis om een leerkracht in staat te stellen leervorderingen te evalueren en vervolgstappen te formuleren. Daarnaast wordt in het document 'Bijlage 2 - Boom LVS Taalverzorging: Spelling Handleiding' in Hoofdstuk 2, paragraaf 2.3, een toelichting gegeven op het analyseren van fouten. Het maken van een uitgebreide foutenanalyse is optioneel. Het analyseprogramma is in staat om alle mogelijke fouten te classificeren en te categoriseren. Dit levert aanvullende informatie op voor de leerkracht. Op bladzijde 43 wordt een voorbeeld van een foutenanalyse getoond en kort besproken.	Ja

### 3. Psychometrische aspecten

#### 3.3.1 Kwaliteit van de dataverzameling

Code	Vragen	Toelichting	Antwoord
P.IV.1	Is de steekproef van leerlingen groot genoeg voor het gekozen model en representatief voor de doelpopulatie?	<p>Het instrument is in inhoud niet veranderd ten opzichte van de vorige indiening. Zodoende heeft de aanbieder afnamedata uit de schooljaren 2023/2024 en 2024/2025 gebruikt als pretestdata – zoals beschreven in paragraaf 4.2 van 'Bijlage 2 - Boom LVS Taalverzorging Spelling Handleiding'. Uit deze afnamedata is een gestratificeerde steekproef getrokken met de stratificatievariabelen schoolweging, schoolgrootte, regio en stedelijkheid, en worden sbo-leerlingen buiten beschouwing gelaten (conform vereist in het beoordelingskader). Hiervoor zijn CBS-gegevens gebruikt. De verschillende fasen in de steekproefprocedure worden beschreven in paragraaf 4.2.1.</p> <p>Voor de blokken van versies A en B geldt dat voor iedere combinatie van blokken een aparte steekproef is getrokken van 1000 afnames (zie Tabellen 4.1 en 4.2). Daarmee wordt voldaan aan de eis die stelt dat er minstens 300 observaties per item dienen te zijn voor het schatten van een 1PL-model.</p> <p>Aangezien deze pretestdata afkomstig zijn uit operationele afnames, zijn de afnamecondities pretest/operationeel identiek.</p> <p>De aanbieder laat zien hoe de steekproef zich verhoudt tot de landelijke populatiewaarden (Tabel 4.3 t/m 4.10, 'aanvullende informatie'). Hier wordt een Cohen's W bij gepresenteerd als effect size om de verschillen tussen steekproef en populatie te kunnen duiden.</p> <p>Uit de resultaten blijkt dat de steekproef voldoende representatief is voor de landelijke populatie (verwaarloosbare tot kleine effecten).</p>	Ja

		Daarmee beoordeelt het beoordelingsteam deze eis als voldoende.	
P.IV.2	Indien er sprake is van een onvolledig dataverzamelingsdesign: is het dataverzamelingsdesign adequaat?	<p><b>Design</b></p> <p>In paragraaf 4.4 van 'Bijlage 2 - Boom LVS Taalverzorging Spelling Handleiding' worden de designs weergegeven voor Spelling Woorden (Tabel 4.11) en Spelling Werkwoorden (Tabel 4.12).</p> <p>Voor Spelling Woorden zijn er zes blokken voor zowel versie A als B. De blokken lopen in zowel versie A als B op in moeilijkheid en ze worden bij een afname altijd gecombineerd aangeboden, waardoor er een gelinkt design is met vijf toetsboekjes per versie (12A, 23A, 34A, 45A, 56A en 12B, 23B, 34B, 45B, 56B). Toetsboekje 12A (of B12) is het makkelijkst en 56A (of 56B) het moeilijkst. Ieder blok bestaat uit 15 items, dus elk toetsboekje bestaat uit 30 items.</p> <p>Voor Spelling Werkwoorden zijn er drie blokken voor zowel versie A als B. De blokken lopen in zowel versie A als B op in moeilijkheid en ze worden bij een afname altijd gecombineerd aangeboden, waardoor er een gelinkt design is met twee toetsboekjes per versie (12A, 23A en 12B, 23B). Toetsboekje 12A (of B12) is het makkelijkst en 23A (of 23B) het moeilijkst. Ieder blok bestaat uit 18 items, dus elk toetsboekje bestaat uit 36 items.</p> <p>In de designs is af te lezen welke toetsen in beginsel zijn bedoeld voor welke deelpopulaties, i.e., leerlingen van een bepaalde DL (didactische leeftijd).</p> <p>Hiermee heeft de aanbieder voldoende beschreven hoe het design eruitziet: het totaal aantal boekjes, hoe de boekjes worden toegewezen aan welke leerlingen en het aantal items per blok/toetsversie. Het beoordelingsteam acht de link tussen de toetsboekjes voldoende stevig.</p>	Ja

		<p><b>Kalibratie</b></p> <p>De aanbieder past een 1PL-model (met JML-schatter) op de data. De kalibratie is als volgt uitgevoerd:</p> <ul style="list-style-type: none"><li>- Eerst zijn de itemparameters van toetsversie 12A geschat. Hierbij zijn vijf items gefixeerd op de reeds gebruikte parameters. De andere items van 12A zijn vrij geschat.</li><li>- Vervolgens worden de itemparameters van toetsversie 23A geschat waarbij de items van blok 2 zijn gefixeerd (doordat die al zijn geschat in blok 12A).</li><li>- Dit herhaalt zich tot en met de laatste toetsversie.</li></ul> <p>De vaardigheidsschattingen zijn gedaan met ML. Het beoordelingsteam had in eerste instantie bedenkingen bij de het gebruik van de JML-methode en de sequentiële kalibraties die de aanbieder had uitgevoerd, in tegenstelling tot bijvoorbeeld het uitvoeren van een concurrente kalibratie met CML. De aanbieder kon aantonen dat het in de praktijk geen verschil maakt welke van de twee methoden werd gebruikt. Daarmee is de aanbieder voldoende overtuigd van de gebruikte methode.</p> <p>De parallelliteit van versies A en B wordt besproken in paragraaf 4.5. Versies A en B zijn op dezelfde schaal gezet doordat zowel voor versie A als B 5 items zijn gefixeerd op de oude, reeds gebruikte parameters, die op dezelfde schaal staan. Het beoordelingsteam ziet dat de informatiecurves van de parallelle versies haast niet te onderscheiden zijn van elkaar. Daarmee is de aanbieder voldoende overtuigd van de gebruikte methode.</p> <p><b>Verschillen itemparameters tussen toetsversies</b></p> <p>Voor sommige items gold dat de geschatte itemparameters van het ene blok niet pasten op de empirische data van het andere blok. In die gevallen is de fixatie losgelaten en zijn itemparameters apart geschat per blok. In Tabel</p>	
--	--	--	--

		<p>4.14 wordt aangegeven om welke items het gaat. Het beoordelingsteam kan niet nagaan in welke mate de items uit Tabel 4.14 misfit vertoonden, maar het beoordelingsteam stelt dat het verstandig is geweest om items los te koppelen bij een misfit. Er zijn dan ook geen opvallendheden in de ICC's van de betreffende items en de link tussen de toetsversies blijft voldoende sterk.</p> <p><b>DIF schoolweging</b></p> <p>In paragraaf 4.4 van 'Boom LVS Taalverzorging Spelling Handleiding' wordt een onderzoek naar DIF ten aanzien van schoolweging (benedengemiddeld versus bovengemiddeld) beschreven. Hier worden Mantel-Haenszeltoetsen op itemniveau uitgevoerd en de aanbieder heeft daarbij de ETS Delta effect size (uit het difR R-package) gehanteerd ter evaluatie.</p> <p>Uit de DIF-analyses blijkt dat er een aantal items zijn met Delta ETS van C (groot effect; twee items bij Spelling Woorden en vier items bij Spelling Werkwoorden). Hierbij geeft de aanbieder aan dat het enkel voorkwam in één van de twee toetsversies waarin het woord voorkwam. Als het voorkwam bij de ene toetsversie, dan kwam het niet voor bij de andere toetsversie. In de aanvullende informatie wordt voor de woorden een mogelijke verklaring gegeven. De items blijven gehandhaafd en het beoordelingsteam vindt het gegeven het lage aantal items (twee en vier van de respectievelijk 300 en 144 vergelijkingen bij Spelling Woorden en Spelling Werkwoorden) en het gebrek aan een duidelijk te verklaren oorzaak gerechtvaardigd.</p> <p><b>Modelpassing en grafische weergave items</b></p> <p>De modelpassing is geëvalueerd door voor ieder item te berekenen of de geschatte p-waarde en de geobserveerde p-waarde niet meer dan 0,1 verschil laten zien. De resultaten hiervan zijn weergegeven in Bijlage XIII en IX. Geen enkel item liet een afwijking zien – na loslaten van de fixatie van de</p>	
--	--	--	--

		<p>items uit Tabel 4.14. In Bijlage X en XI worden ICC's weergegeven van ieder item in ieder toetsboekje. Hieruit blijkt volgens het beoordelingsteam dat het model voldoende goed past bij de data.</p> <p><b>Scoreverdeling en betrouwbaarheid</b></p> <p>De aanbieder levert de scoreverdeling aan voor ieder toetsboekje en geeft daarbij aan voor welke scores geldt dat de betrouwbaarheid van de meting hoger is dan 0,7. Enkel bij de meest extremere scores valt – per definitie - de betrouwbaarheid lager uit, hetgeen het beoordelingsteam acceptabel acht.</p> <p><b>Vermoeidheidseffecten</b></p> <p>Zowel Spelling Woorden als Spelling Werkwoorden wordt afgenomen als een dictee. Ze geeft aan dat dit in de praktijk betekent dat de toetsen binnen een half uur kunnen worden afgenomen. Mocht zich inderdaad vermoeidheidseffecten voordoen, bijvoorbeeld bij de lagere groepen, dan is er de mogelijkheid de twee blokken van de toets op verschillende momenten af te nemen. Het beoordelingsteam is hiermee akkoord.</p> <p><b>Conclusie</b></p> <p>Daarmee beoordeelt het beoordelingsteam deze eis als voldoende.</p>	
--	--	--	--

### 3.3.2 Normering

#### Bepalen van absolute normen

Code	Vragen	Toelichting	Antwoord
P.V.1	Is de keuze voor de standaardbepalingsmethode gedegen gemotiveerd en op de juiste wijze uitgevoerd?		N.v.t.

Code	Vragen	Toelichting	Antwoord
P.V.2	Zijn de experts met inhoudelijke vakdeskundigheid die de standaarden (cesuren) vaststellen naar behoren geselecteerd en getraind?		N.v.t.
P.V.3	Is er voldoende overeenstemming tussen de experts?		N.v.t.

#### Vaststellen van relatieve normen

Code	Vragen	Toelichting	Antwoord
P.V.4	Als de normeringssteekproef afwijkt van de kalibratiesteekproef, zijn de normgroepen groot genoeg?	In Tabel 5.1 en 5.2 van 'Bijlage 2 - Boom LVS Taalverzorging Spelling Handleiding' wordt de omvang van de normeringssteekproef weergegeven. Het blok met het laagste aantal observaties heeft 1325 observaties, hetgeen ruim meer is dan de conform het kader vereiste minimumaantal van 300 observaties per item.  Daarmee beoordeelt het beoordelingsteam deze eis als voldoende.	Ja
P.V.5	Als de normeringssteekproef afwijkt van de kalibratiesteekproef, zijn de normgroepen representatief?	De aanbieder geeft een beschrijving van de samenstelling van de normeringssteekgroep voor de achtergrondvariabelen regio, stedelijkheid, schoolgrootte en schoolweging in paragraaf 5.3 van 'Boom LVS Taalverzorging Spelling Handleiding'. Ze heeft op deze variabelen een aselechte steekproef getrokken, op eenzelfde manier als bij de kalibratiesteekproef (zie eis P.IV.1). Net als bij de kalibratiesteekproef wordt gebruikgemaakt van data van schooljaren	Ja

		<p>2023-2024 en 2024-2025. Het verschil met de kalibratiesteekproeftrekking was dat:</p> <ul style="list-style-type: none"><li>- Ook gebruikgemaakt is van toetscores waarvan alleen de totaalscore werd geregistreerd, terwijl de kalibratiesteekproef bestond uit leerlingen waarvan de data op itemniveau geregistreerd was;</li><li>- De data werden geselecteerd uit specifieke afnamemaanden (voornamelijk januari en juni, zie paragraaf 5.2 en Tabel 5.1 voor een beschrijving)</li></ul> <p>De aanbieder levert Cohen's W aan om de verdeling van de steekproef te vergelijken met de populatieverdeling. Er worden geen grote effecten gevonden per toets (<math>W &lt; 0,5</math>).</p> <p>Ten aanzien van de variabele 'regio' zijn de afwijkingen het grootst bij Spelling Woorden, waarbij Oost en Noord oververtegenwoordigd zijn en waarbij Zuid en West ondervertegenwoordigd zijn. Voor schoolweging geldt dat aan de hand van Cohen's W middelgrote tot grote effecten worden gevonden. Hier geldt echter voor dat de percentages voor de uiterste categorieën in de landelijke verdeling erg laag zijn. Kleine afwijkingen hiervan worden al snel opgevat als groot. De aanbieder berekent daartoe aanvullend Cohen's W voor schoolweging met als aanpassing dat de twee uiterste categorieën 1 en 5 opgaan in, respectievelijk, 2 en 4. Na aanpassing worden enkel verwaarloosbare tot kleine effecten gevonden.</p> <p>Het beoordelingsteam vindt dit voldoende overtuigend en daarmee beoordeelt het beoordelingsteam deze eis als voldoende.</p>	
--	--	--	--

P.V.6	Zijn de normen correct bepaald?	<p>In paragraaf 5.4 van 'Boom LVS Taalverzorging Spelling Handleiding' wordt de totstandkoming van de normen beschreven.</p> <p>Zoals bij P.V.5 beschreven zijn data gebruikt van specifieke afnamemaanden. Daarmee zijn er dus data van specifieke didactische leeftijden (DL, en zie Tabel 5.1 voor welke DL voor welke toetsboekjes zijn gebruikt in de normeringssteekproef).</p> <p>Voor iedere DL van een toetsboekje zijn Z-scores geschat als functie van de vaardigheidsscore, de DL en de interacties tussen beide variabelen (de normeringsfunctie). Voor de tussenliggende DL's (DL's waarop niet genormeerd is), zijn Z-scores geïnterpoleerd, waarbij uitgegaan wordt van een continue groei in de populatie.</p> <p>Tot slot zijn de empirische Z-scores omgezet naar percentielscores. De percentielscores zijn gecategoriseerd tot I-V- en A-E-scores. Ook wordt een DLE (didactische leeftijdsequivalent) gegeven: een DL op basis van de behaalde score (zie ook beantwoording eis P.VIII.1).</p> <p>Het beoordelingsteam veronderstelt dat de normen correct zijn bepaald en daarmee beoordeelt het beoordelingsteam deze eis als voldoende.</p>	Ja
-------	---------------------------------	--	----

[Linken met referentieniveaus](#)

Code	Vragen	Toelichting	Antwoord
P.V.7	Is de toets op gedegen manier verbonden aan passende referentieniveaus?		N.v.t.

### Andere methodes

Code	Vragen	Toelichting	Antwoord
P.V.8	Als naast/in plaats van de bovengenoemde methoden (eisen P.V.1 t/m P.V.7) ook andere normeringsmethoden worden gebruikt, zijn deze methodes goed omschreven, voldoende onderbouwd en correct uitgevoerd?	Er is sprake van continu normeren. Dit is reeds besproken bij P.V.4 t/m P.V.6, aangezien het hier ook een vorm van relatief normeren betreft.	N.v.t.

### 3.3.3 Betrouwbaarheid

Code	Vragen	Toelichting	Antwoord
P.VI.1	Zijn of worden de betrouwbaarheidsgegevens van de metingen correct berekend?	<p>In paragraaf 6.2 van 'Bijlage 2 - Boom LVS Taalverzorging Spelling Handleiding' wordt aangegeven dat de globale betrouwbaarheid wordt berekend aan de hand van Cronbach's Alpha en McDonald's Omega.</p> <p>In paragraaf 6.3 beschrijft de aanbieder dat ze de lokale betrouwbaarheid heeft geëvalueerd door een meetnauwkeurigheid van 0,7 af te zetten tegen de verdeling informatiewaarden per toetsboekje (zie Bijlage IV en V, aanvullende informatie). Hierdoor is voor het beoordelingsteam duidelijk te zien vanaf en tot welke score op de verschillende toetsen voldoende nauwkeurig wordt gemeten.</p> <p>Daarmee beoordeelt het beoordelingsteam deze eis als voldoende.</p>	Ja
P.VI.2	Zijn de betrouwbaarheidsgegevens voldoende gezien de conclusies en	In Tabel 6.1 t/m 6.4 van 'Bijlage 2 - Boom LVS Taalverzorging Spelling Handleiding' worden de betrouwbaarheidsindices voor de globale betrouwbaarheid (Cronbach's Alpha en McDonalds Omega)	Ja

	eventuele beslissingen die met het instrument genomen worden?	<p>weergegeven voor ieder toetsboekje. Deze zijn alle hoger dan de in het kader vereiste 0,7. Alle toetsen zijn uitgevoerd op &gt; 300 leerlingen.</p> <p>In Bijlage XVI en XVII worden de informatiefuncties van de toetsboekjes gepresenteerd. Hierin is te zien dat bij enkel de extremere scores – per definitie - de meetprecisie van 0,7 niet wordt behaald.</p> <p>Het beoordelingsteam vindt dit voldoende overtuigend en daarmee beoordeelt het beoordelingsteam deze eis als voldoende.</p>	
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)?	<p>De aanbieder hanteert de methode van Rudner (2001, 2005) waarbij gekeken naar de kans op correcte/incorrecte classificatie van een behaald niveau (I-V en A-E), gegeven de vaardigheidsscores en SE van de leerlingen in relatie tot de niveauscoringen van niveau I-V en A-E, i.e., de leerling een bepaald niveau wordt toegeschreven gegeven de vaardigheidsscore. Wat is – gegeven de SE - de kans dat deze classificatie terecht is? Zie paragraaf 6.4 van 'Bijlage 2 - Boom LVS Taalverzorging Spelling Handleiding' voor een beschrijving. Deze procedure is gehanteerd voor ieder toetsboekje. De resultaten worden weergegeven in Bijlage XX en XXI. De classificatie wordt gedefinieerd aan de hand van de accuratesse en de plus-minus-1-regel, zoals het beoordelingskader vereist. De plus-minus-1-regel is tussen de 95% en 99% voor de verschillende toetsen. De accuratesse voor de verschillende toetsen zit tussen de 0,58 tot 0,71. De aanbieder haalt literatuur aan waarbij wordt aangegeven dat de waarden in de praktijk tussen 0,6 en 0,7 liggen. Het beoordelingsteam vindt de behaalde waarden acceptabel. Slechts in enkele gevallen is de accuratesse &lt; 0,6 (maar niet lager dan 0,58) en is daarmee van</p>	Ja

		<p>mening dat voldoende is aangetoond dat het instrument accurate metingen doet.</p> <p>Daarmee beoordeelt het beoordelingsteam deze eis als voldoende.</p>	
--	--	---	--

### 3.3.4 Constructvaliditeit

Code	Vragen	Toelichting	Antwoord
P.VII.1	Is aangetoond dat er sprake is van unidimensionaliteit?	<p>In paragraaf 4.3.2 van 'Bijlage 2 - Boom LVS Taalverzorging Spelling Handleiding' beschrijft de aanbieder dat ze een factoranalyse heeft uitgevoerd op de tetrachorische correlaties. Aan de hand van de vertrouwenwekkende RMSEA-resultaten heeft de aanbieder voldoende aannemelijk gemaakt dat er sprake is van unidimensionaliteit.</p> <p>Daarmee beoordeelt het beoordelingsteam deze eis als voldoende.</p>	Ja
P.VII.2	Is aangetoond dat de itemkwaliteit op orde is?	<p>In paragraaf 7.2 van het document 'Bijlage 2 - Boom LVS Taalverzorging Spelling Handleiding' presenteert de aanbieder klassieke itemstatistieken per toetsboekje voor verschillende normeringsmaanden in de normeringssteekproef. Hier worden per toetsboekje de minimale en maximale p- en rir-waarde gegeven. In Bijlage VI en VII van de aanvullende informatie worden staafdiagrammen op itemniveau weergegeven waarbij iedere p- en rir-waarde te inspecteren is.</p> <p>De aanbieder geeft een range van 0,20 – 0,85 waarbinnen de p-waarden idealiter vallen. Als ondergrens voor de rir hanteert de aanbieder 0,1. Ze geeft daarbij aan dat het hier gaat om streefwaarden. Enkele afwijkingen hiervan vindt de aanbieder toegestaan. Het beoordelingsteam gaat hierin mee, temeer omdat het</p>	Ja

		<p>hier een eigen regel betreft en niet een voorgeschreven regel uit het beoordelingskader. Daarnaast onderschrijven de staafdiagrammen dat afwijkingen inderdaad slechts bij enkele items voorkomen.</p> <p>Daarmee beoordeelt het beoordelingsteam deze eis als voldoende.</p>	
P.VII.3	<p>Is de convergente en divergente validiteit van het instrument door onderzoek aangetoond?</p>	<p>In paragraaf 7.3 van 'Bijlage 2 - Boom LVS Taalverzorging Spelling Handleiding' beschrijft de aanbieder hoe ze onderzoek heeft gedaan naar de convergente en divergente validiteit van zowel Spelling woorden als Spelling werkwoorden. Hiertoe heeft ze correlaties berekend tussen de prestaties op Spelling Woorden/Werkwoorden en de prestaties op andere toetsen van de aanbieder.</p> <p>Voor de evaluatie van de <u>convergente validiteit</u> zijn de correlaties berekend tussen:</p> <ul style="list-style-type: none"> <li>- Spelling Woorden en Spelling Werkwoorden</li> </ul> <p>Voor de evaluatie van de <u>divergente validiteit</u> zijn de correlaties berekend tussen:</p> <ul style="list-style-type: none"> <li>- Spelling Woorden/Werkwoorden en Begrijpend Lezen</li> <li>- Spelling Woorden/Werkwoorden en Rekenen-Wiskunde</li> </ul> <p>In Tabel 7.5 worden de correlaties weergegeven tussen de prestaties op Spelling Woorden en Spelling Werkwoorden</p> <p>In Tabel 7.6 en 7.8 worden de correlaties weergegeven tussen respectievelijk Spelling Woorden en andere domeinen, en Spelling Werkwoorden en andere domeinen.</p> <p>Uit de tabellen volgt dat de correlaties tussen Spelling Woorden en Spelling Werkwoorden (marginaal) hoger zijn dan tussen de twee domeinen en andere domeinen.</p> <p>Daarmee beoordeelt het beoordelingsteam deze eis als voldoende.</p>	Ja

P.VII.4	<p>Is aangetoond dat er geen sprake is van item bias voor relevante subgroepen/categorieën?</p>	<p>In paragraaf 7.5 van 'Boom LVS Taalverzorging Spelling Handleiding' wordt een onderzoek naar itembias (DIF) ten aanzien van stedelijkheid (platteland versus verstedelijkt) beschreven. Hier worden Mantel-Haenszeltoetsen op itemniveau uitgevoerd.</p> <p>De aanbieder heeft Mantel-Haenszeltoetsen op itemniveau uitgevoerd en heeft daarbij de ETS Delta effect size (uit het difR R-package) gehanteerd ter evaluatie.</p> <p>Uit de DIF-analyses blijkt dat er een aantal items zijn met Delta ETS van C (groot effect). Hierbij geeft de aanbieder aan dat dit – met uitzondering van twee woorden bij Spelling Woorden – enkel voorkwam in één van de twee toetsversies waarin het woord voorkwam. Als het voorkwam bij de ene toetsversie, dan kwam het niet voor bij de andere toetsversie. Daarnaast geeft de aanbieder ook aan dat de groepen van de categorie 'niet verstedelijkt' in sommige gevallen klein uitvielen, wat interpretatie van de resultaten bemoeilijkt. Ook geeft ze aan dat in veel gevallen geen logische verklaring te vinden valt voor de gevonden DIF. In de aanvullende informatie wordt voor enkele woorden een mogelijke verklaring gegeven.</p> <p>Al met al heeft bovenstaande de aanbieder doen besluiten om items niet te verwijderen op basis van de gedetecteerde DIF. De aanbieder gaat wel een aantal contextzinnen veranderen bij de items met gevonden DIF, en zal in de toekomst monitoren of deze aanpassing leidt tot minder bias (zie ook beantwoording eis O.IV.2).</p> <p>Het beoordelingsteam gaat mee in de redentie van de aanbieder en daarmee wordt deze eis als voldoende beoordeeld.</p>	Ja
---------	---	--	----

P.VII.5	Zijn verschillen in presentaties van relevante subgroepen voldoende onderbouwd?	<p>De aanbieder heeft in paragraaf 7.4 van 'Bijlage 2 - Boom LVS Taalverzorging Spelling Handleiding' de prestaties vergeleken tussen subgroepen van de achtergrondvariabelen Regio en Schoolweging. De subgroepen zijn gedefinieerd zoals wordt beschreven in het beoordelingskader bij eis P.IV.1. De aanbieder hanteert een <math>\eta^2</math> voor het evalueren van de groepsverschillen, wat een relevante statistiek is. Ze vindt voornamelijk kleine/verwaarloosbare verschillen voor Regio, alleen valt het op dat regio West slechter presteert in de onderbouw op Spelling Woorden vergeleken met de andere regio's (<math>\eta^2 = 0,05</math>; middelgroot effect = <math>0,06</math>). Ze schrijft dit toe aan het feit dat schoolweging ook lager is in West.</p> <p>Voor de variabele Schoolweging wordt een klein tot middelgroot effect gevonden (<math>\eta^2 = 0,04</math>) bij Spelling Werkwoorden bij met name eind groep 7. In Tabel 7.13 is af te lezen dat de leerlingen uit scholen met lagere schoolweging hoger presteren. Hiervoor presenteert de aanbieder enkele mogelijke verklaringen (leerlingen van scholen met een laag schoolgewicht zijn afkomstig uit gezinnen met een relatief hoge sociaaleconomische status. In die gezinnen worden kinderen vaak ook meer voorgelezen en meer gestimuleerd om zelf te lezen).</p> <p>Hiermee heeft de aanbieder een duiding gegeven van het verschil en daarmee beoordeelt het beoordelingsteam deze eis als voldoende.</p>	Ja
---------	---	--	----

### 3.3.5 Volgaspect

Code	Vragen	Toelichting	Antwoord
P.VIII.1	<p>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>	<p>De aanbieder heeft aangetoond dat de groei adequaat wordt gemeten. In paragraaf 3.4 en 3.6 van 'Bijlage 2 - Boom LVS Taalverzorging Spelling Handleiding' zet de aanbieder uiteen hoe betekenis wordt gegeven aan het concept 'volgsysteem'.</p> <p>Zoals in paragraaf 4.4 beschreven, geldt voor beide instrumenten (Spelling Woorden en Spelling Werkwoorden) dat er sprake is van een gelinkt design (zie ook beantwoording eis P.IV.2), waardoor de scores van de leerlingen op deze toetsen op één schaal staan (één schaal voor Spelling Woorden en één schaal voor Spelling Werkwoorden). Hierdoor kunnen de prestaties op de verschillende meetmomenten met elkaar vergeleken worden, en zo kan de vaardigheid van de leerling in kaart worden gebracht door de tijd heen.</p> <p>Bij iedere didactische leeftijd (DL) hoort een verdeling van vaardigheidsscores (die volgen uit de normeringssteekproef). De vaardigheidsscore van een leerling van een bepaalde DL wordt na een meetmoment afgezet tegen de verdeling vaardigheidsscores van de betreffende DL. Zo kan per meetmoment worden geëvalueerd of er groei is en of de groei volgens verwachting is, dan wel sneller of langzamer dan verwacht. Dit wordt uitgedrukt in De Didactische Leeftijdsequivalent (DLE), i.e., indien een leerling bijvoorbeeld hoger scoort op een toets dan gemiddeld voor diens DL, wordt beschouwd voor welke DL het de gemiddelde score betreft. De DLE ligt in een</p>	Ja

		<p>dergelijk geval hoger dan de DL (zie paragraaf 3.4 voor een expliciete beschrijving).</p> <p>De aanbieder zet in paragraaf 3.4 t/m 3.6 duidelijk uiteen voor gebruikers hoe de resultaten geïnterpreteerd dienen te worden.</p> <p>Daarmee beoordeelt het beoordelingsteam deze eis als voldoende.</p>	
P.VIII.2	Wordt de groei van een leerling nauwkeurig en betrouwbaar gemeten?	<p>Om te bepalen wanneer bij een hogere of lagere score sprake is van 'echte verandering' heeft de aanbieder gekozen voor het berekenen van de meetprecisie van de verschillcores aan de hand van de Reliable Change Index (RCI; zie aanvullende informatie). Hierbij hanteert de aanbieder een 90%-betrouwbaarheidsinterval. De aanbieder geeft in Bijlagen XVI en XVII visueel weer bij welke scores op welke twee scoremomenten sprake is van achteruitgang of vooruitgang, of dat dat niet te zeggen is gegeven het 90%-betrouwbaarheidsinterval. Hiermee is het voor het beoordelingsteam inzichtelijk bij welke verschillcores sprake is van verandering die niet toe te schrijven is aan meetfout.</p> <p>Daarmee beoordeelt het beoordelingsteam deze eis als voldoende.</p>	Ja
P.VIII.3	Worden er gegevens verstrekt over hoe groei geïnterpreteerd dient te worden?	<p>De aanbieder zet in paragraaf 3.4 t/m 3.6 van 'Bijlage 2 - Boom LVS Taalverzorging Spelling Handleiding' duidelijk uiteen voor gebruikers hoe de resultaten geïnterpreteerd dienen te worden. Daarnaast geeft de ouderbrochure ('Bijlage 3 - Boom LVS ouderbrochure') duidelijke tekst en uitleg bij een plot waar meerdere meetmomenten van een leerling worden weergegeven (met betrouwbaarheidsinterval), ten opzichte van verwachte scores gegeven de didactische leeftijd van de leerling op de verschillende meetmomenten.</p> <p>Daarmee beoordeelt het beoordelingsteam deze eis als voldoende.</p>	Ja

#### 4. Organisatorische aspecten

Code	Vragen	Toelichting	Antwoord
B.I.1	Heeft de aanbieder een risico-analyse omtrent informatiebeveiliging en privacy uitgevoerd bij het gebruik van het LVS-instrument?	De aanbieder is ISO 27001 gecertificeerd. De kwaliteitseisen behorende bij de organisatorische aspecten zijn hierdoor voldoende verantwoord.	Ja
B.I.2	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?	De aanbieder is ISO 27001 gecertificeerd. De kwaliteitseisen behorende bij de organisatorische aspecten zijn hierdoor voldoende verantwoord.	Ja
B.I.3	Hebben de functionarissen van de aanbieder die betrokken zijn bij de inhoud en de beveiliging van het LVS-instrument aantoonbaar een geheimhoudingsverklaring ondertekend?	De aanbieder is ISO 27001 gecertificeerd. De kwaliteitseisen behorende bij de organisatorische aspecten zijn hierdoor voldoende verantwoord.	Ja
B.I.4	Waarborgt de aanbieder op gedegen en complete wijze de privacy van leerkracht, leerling en ouder conform de Algemene Verordening Gegevensbescherming?	De aanbieder is ISO 27001 gecertificeerd. De kwaliteitseisen behorende bij de organisatorische aspecten zijn hierdoor voldoende verantwoord.	Ja
B.I.5	Hanteert de aanbieder op gedegen en complete wijze incidentmanagement inclusief een protocol datalekken, zoals voorgeschreven door de Autoriteit Persoonsgegevens?	De aanbieder is ISO 27001 gecertificeerd. De kwaliteitseisen behorende bij de organisatorische aspecten zijn hierdoor voldoende verantwoord.	Ja

Code	Vragen	Toelichting	Antwoord
B.I.6	Heeft de aanbieder duidelijke richtlijnen (in het toetsreglement/ verantwoordingsdocumenten) opgesteld voor bij de afname betrokkene personen (zoals de schoolleider en leerkracht) over de veilige opslag en het veilige gebruik van het instrument en de daarmee verzamelde gegevens?	De aanbieder is ISO 27001 gecertificeerd. De kwaliteitseisen behorende bij de organisatorische aspecten zijn hierdoor voldoende verantwoord.	Ja
B.I.7	Heeft de aanbieder een protocol (in het toetsreglement / verantwoordingsdocumenten) opgesteld voor een regelmatige controle of de inhoud van het LVS-instrument of andere niet-openbare informatie over het LVS-instrument uitgelekt/openbaar geworden is, bijv. door regelmatige controle van het internet en andere media?	De aanbieder is ISO 27001 gecertificeerd. De kwaliteitseisen behorende bij de organisatorische aspecten zijn hierdoor voldoende verantwoord.	Ja