

1. Algemene informatie

Algemeen en meetpretentie

Het IEP LVS is een methodeonafhankelijk volgsysteem waarin leerlingen vanaf leerjaar 3 tot aan de eindtoets gevolgd kunnen worden in hun ontwikkeling. Het IEP-LVS is een leer- en criteriumgericht volgsysteem.

De toetsen Technisch Lezen 6a en 6b hebben hetzelfde doel en functie als de toetsen Technisch Lezen 3, 4 en 5, die in 2021 reeds zijn goedgekeurd door de Expertgroep. Tezamen vormen de toetsen één integraal leerlingvolgsysteem voor de leerjaren 3 t/m 6.

Met de IEP-volgttoetsen taal en rekenen wordt gemeten op welk niveau de leerlingen de onderdelen Lezen, Technisch Lezen, Taalverzorging en Rekenen beheersen. Het inhoudelijk toetskader (zie hoofdstuk 2) vormt de basis voor de inhoud van de toetsen Technisch Lezen leerjaar 6 en is verdeeld in twee niveaubeschrijvingen (6a en 6b).

Doelgroep

Het IEP LVS bevat toetsen voor leerlingen in leerjaar 3 tot en met 8. Het IEP LVS is zo ingericht dat alle toetsen voor elke leerling toegankelijk zijn. Zo kan een leerling in leerjaar 5 die moeite heeft met taalverzorging, ook de toets op het niveau van leerjaar 4 maken. Maar het is bijvoorbeeld ook mogelijk om een leerling in leerjaar 6 die relatief goed is in rekenen al een toets op 2F-niveau te geven. Doordat het mogelijk is om bij een leerling toetsen van verschillende niveaus af te nemen, kun je goed aansluiten bij het individuele niveau van de leerling. Dit maakt dat het IEP LVS ook geschikt is voor leerlingen uit het SBO.

Inhoudelijke theoretische inkadering:

De IEP LVS toetsen voor de leerjaren 6, 7 en 8 zijn gebaseerd op het Referentiekader taal en rekenen (Meijerink et al., 2009). Deze toetsen geven aan of de leerling het gemeten referentieniveau beheerst.

Met de IEP-volgttoetsen taal en rekenen wordt gemeten op welk niveau de leerlingen de onderdelen Lezen, Technisch Lezen, Taalverzorging en Rekenen beheersen. Het inhoudelijk toetskader vormt de basis voor de inhoud van de toetsen Technisch Lezen leerjaar 6 en is verdeeld in twee niveaubeschrijvingen (6a en 6b)

Inhoud van het toetspakket:

Het toetspakket Technisch Lezen 6a6b bestaat uit de volgende documenten:

- Verantwoording LVS "Technisch Lezen" groep 6a en 6b; deze bevat informatie over:
 - o De uitgangspunten van de toetsen (hfdst. 2),
 - o De inhoud van de toetsen (hfdst. 3),
 - o Normeringspopulatie (hfdst. 4),
 - o Het design van de dataverzameling (hfdst. 5),
 - o De kalibratie en kwaliteit van de items (hfdst. 6),
 - o De bepaling van cesuren (hfdst. 7),
 - o De constructvaliditeit (hfdst. 8),
 - o Het volgaspect (hfdst. 9), en
 - o Inzicht in de leervorderingen (hfdst. 10).

Daarnaast werden de volgende bijlagen beschikbaar gesteld:

- Toetswijzer (bijlage 1)
- Boekje leerkracht Technisch Lezen 6a (bijlage 2a)
- Boekje leerkracht Technisch Lezen 6b (bijlage 2b)
- Itemoverzicht (bijlage 2c)
- Toelichtingen (bijlage 3)
- Itemparameters (bijlage 4)
- TIA's (bijlage 5)
- DIF (bijlage 6)
- Lagrange Multiplier Tracelines (bijlage 7)
- Algemene toelichting methode (bijlage 8)
- Handreiking interpreteren toetsresultaten (bijlage 9)
- Talentenkaart (bijlage 10)
- Leeswijzer versie leerkracht (bijlage 11)

2. Beoordeling van de kwaliteitsaspecten

De beoordeling vindt plaats volgens het 'Beoordelingskader voor instrumenten binnen leerlingvolgsystemen (LVS)', zoals opgesteld door de Expertgroep Toetsen PO. De Expertgroep Toetsen PO wordt gevormd door Prof. Dr. Cees Van der Vleuten (voorzitter), Prof. dr. Cees Glas (psychometrisch expert), Prof. Dr. Desiree Joosten-Ten Brinke (onderwijskundig expert) en Liza Kozłowska MA (secretaris).

Bij onderstaande beoordeling van de kwaliteitsaspecten met bijbehorende codes van het voornoemde beoordelingskader worden passages uit de wetenschappelijke verantwoording en de Handleiding veelal letterlijk vermeld.

De kwaliteit van de dataverzameling

S1 Is de steekproef van leerlingen representatief?

Bevindingen:

Uitgangspunt is een steekproef die resulteert in tenminste 400 observaties per item. Hoewel voor de kalibratie van de toetsen met het hier gehanteerde 1PL-model (Rasch model) het minimale aantal observaties 200 is, is voor het evalueren van de modelpassing van een redelijk alternatief model, hier het 2PL-model, minimaal 400 de norm. Uit bijlage 4 (zie kolom 'Aantal afnames') blijkt dat het aantal observaties van de IEP LVS items conform de eisen is.

Naast aantallen afnames is de representativiteit op achtergrondgegevens van de normeringspopulatie (groep leerlingen in leerjaar 6 die aan het normeringsonderzoek heeft deelgenomen) ten opzichte van de doelpopulatie (alle leerlingen in leerjaar 6 van het regulier basisonderwijs) van belang om een oordeel te kunnen toekennen aan dit aspect. De auteurs stellen dat voor de bepaling van de kwaliteit van de items (hoofdstuk 6) en de standaardsetting (hoofdstuk 7) het niet noodzakelijk is dat de normeringspopulatie ook een normpopulatie (i.e., een representatieve steekproef) is.

Wel is het van belang vast te stellen dat de normeringspopulatie geen specifiek selecte groep is van de leerlingen in leerjaar 6 van het basisonderwijs, d.w.z. dat er in ieder geval achtergrondgegevens worden gerepresenteerd in de steekproef. De reden hiervoor is dat in het onderhavige onderzoek absolute normen worden bepaald, waardoor de representativiteit van ondergeschikt belang is. Voor de bepaling van de gemiddelde groeifactor van de IEP LVS populatie (hoofdstuk 9) is volgens de auteurs de representativiteit van de normeringspopulatie echter wel van belang, omdat daar een relatieve norm wordt bepaald.

Persoonlijke achtergrondgegevens van de leerlingen konden niet worden gebruikt vanwege het waarborgen van de privacy (zoals beschreven in de Wet bescherming persoonsgegevens) en daarom zijn in dit normeringsonderzoek alleen de schoolachtergrondgegevens denominatie, urbanisatiegraad, schoolgrootte, regio en schoolweging gebruikt om representativiteit te onderzoeken. Deze schoolachtergrondgegevens zijn openbaar beschikbaar per school bij DUO en het CBS. Het digitale platform IEP LVS is gebruikt om de data voor het normeringsonderzoek te verzamelen. De hiervoor gebruikte toetsen zijn door de betrokken scholen tijdens het reguliere onderwijsproces in schooljaar 2020-2021 op eigen initiatief afgenomen in

leerjaar 6. Er is hier dus sprake van 'purposive sampling' (doelsteekproef), een niet-probabilistische steekproeftechniek. De afnamecondities waren gelijk aan de condities waaronder de IEP LVS toetsen afgenomen gaan worden.

De responsedata is voor alle analyses gefilterd op de passendheid van de toets bij het afnamemoment zodat verwacht kan worden dat de gemeten doelen uit de toets zijn aangeboden in het onderwijs.

Voor de analyse voor het bepalen van de gemiddelde groeifactor (hoofdstuk 9) is naast passendheid, gefilterd op twee criteria. Ten eerste zijn records van SBO-scholen en van scholen van Bonaire uit het databestand verwijderd. Ten tweede zijn records verwijderd waarvoor het toetsresultaat niet in een ontwikkelscore uitgedrukt kan worden (hoofdstuk 7.1).

Tabel 4.1 laat zien hoe de normeringspopulatie zich op achtergrondgegevens verhoudt tot de doelpopulatie voor de analyses ten behoeve van de bepaling van de gemiddelde groeifactor (hoofdstuk 9). Door het filteren op andere criteria dan criteria gebruikt voor de analyses ten behoeve van de gemiddelde groeifactor (hoofdstuk 9) is de normeringspopulatie die gebruikt is voor de kwaliteitsanalyse en de standaardsetting (hoofdstuk 6 en 7) weliswaar niet identiek (iets grotere N), maar is wel vergelijkbaar qua verdeling over de diverse achtergrondvariabelen. In Tabel 4.1 wordt daarom de normeringspopulatie weergegeven als gebruikt in hoofdstuk 9, omdat de representativiteit voor deze relatieve normering van belang is.

Uit de verdeling die in tabel 4.1 getoond wordt, is op te maken dat de normeringspopulatie niet op alle categorieën als representatieve random steekproef van leerjaar 6 van het basisonderwijs beschouwd kan worden. Wel kan uit de verdeling in tabel 4.1 geconcludeerd worden dat alle achtergrondcategorieën vertegenwoordigd zijn in de normeringspopulatie en vergelijkbaar verdeeld zijn ten opzichte van de landelijke populatie in het gehele basisonderwijs. Dit impliceert dat voor de toekomst de gemiddelde groeifactor jaarlijks geëvalueerd wordt aan de hand van representatieve afnamedata en indien nodig wordt aangepast (zie hoofdstuk 9).

De gemiddelde groeifactor in het IEP LVS is een factor die gebaseerd wordt op data van leerlingen uit regulier basisonderwijs. Records van leerlingen uit het SBO worden daarom verwijderd. Ook records van leerlingen van scholen van Bonaire worden verwijderd omdat zij Nederlands als tweede taal hebben. Daarmee is technisch lezen in het Nederlands voor hen een net wat andere vaardigheid dan voor leerlingen in het Nederlandse basisonderwijs.

Records waarvoor het toetsresultaat niet in een ontwikkelscore uitgedrukt wordt, worden verwijderd omdat dit betekent dat de toets te moeilijk was voor de leerling. De leerkracht heeft in dat geval niet een passende toets laten maken door de leerling. Het al dan niet geven van een ontwikkelscore bij een toetsresultaat is bepaald in de standaardsetting. Om die reden wordt er dan geen ontwikkelscore gegeven en kan er ook geen groeifactor berekend worden.

Conclusie:

De procedure voor het samenstellen van de steekproeven is onderbouwd en de omstandigheden en momenten waaronder data is verzameld, zijn vergelijkbaar met de afnamecondities en afnamemomenten waaronder de toetsen worden afgenomen. Op

aspect S1 wordt aan de toetsen IEP LVS Technisch Lezen 6a en 6 b het volgende oordeel toegekend: **'voldoende'**

S2 In geval van een onvolledig dataverzamelingsdesign: is het design adequaat?

Bevindingen:

Het normeringsonderzoek voor het IEP LVS leerjaar 6 bestond uit de operationele afname van de pretest IEP LVS leerjaar 6 toetsen in het schooljaar 2020-2021. De pretest IEP LVS leerjaar 6 toetsen bestonden uit meer items dan de IEP LVS leerjaar 6 toetsen definitief bevatten. Op basis van de toets- en itemanalyses op de responsdata van de afnames in het normeringsonderzoek en het design van de definitieve toetsen is de selectie van de items per toets vastgesteld. De afnamecondities van de pretest IEP LVS leerjaar 6 toetsen waren identiek aan de definitieve IEP LVS leerjaar 6 toetsen. Alle deelnemers aan het normeringsonderzoek in de pretestfase hebben beide onderdelen, leestechiek en leesexpressie, gemaakt.

Bij de beschrijving van het design wordt onderscheid gemaakt tussen de onderdelen leestechiek en leesexpressie.

Bij de selectie van de items voor het onderdeel Leestechiek in de IEP LVS toetsen Technisch Lezen is ten eerste gekeken naar het design van de toetsen waaraan de itemselectie van de toetsen moest voldoen. Dit design is in Tabel 6.1 te zien. Hierin is te zien dat de definitieve samenstelling van iedere toets bestaat uit 30 items van het niveau van de toets, 5 items van het niveau eronder en 5 items van het niveau erboven. Daarnaast is er bij de definitieve samenstelling van de toetsen sprake van enige overlap (minimaal 10 items) tussen de opeenvolgende niveautoetsen.

Ten tweede is gekeken naar de statistieken afkomstig uit de TIA-analyse over de data uit het normeringsonderzoek van het onderdeel Leestechiek. Er is rekening gehouden met de p-waarde (proportie correct) en rit-waarde (item-totaalcorrelatie) van de items. De p- en rit -waarde van de items geselecteerd voor de IEP LVS toetsen zijn in het document TIA's IEP LVS toetsen Technisch Lezen (bijlage 5) per toets weergegeven. De uitgangspunten bij deze selectie waren:

- In de samenstelling moet de inhoudelijke dekking van het IEP LVS Toetskader (zie hoofdstuk 2 van de Verantwoording) gewaarborgd zijn.
- Items hebben een rit-waarde van groter of gelijk aan 0,20, zodat de items als voldoende beoordeeld kunnen worden volgens het beoordelingssysteem van de COTAN (Evers, Lucassen, Meijer, & Sijtsma, 2010).
- Items hebben een p-waarde van minimaal 0,50 en maximaal 0,95 of items hebben een p-waarde groter dan 0,95 maar hebben daarbij een rit-waarde van minimaal 0,30.

Voor de IRT-analyse van de IEP LVS toetsen is een one-parameter logistisch model (1PLM, oftewel Rasch model) met de marginal maximum likelihood (MML) schattingsmethode toegepast. De IRT-analyses zijn uitgevoerd in Versie 0.4.0 van de applicatie Lexter (2021). De keuze voor een 1PLM analyse in Lexter is gemaakt omdat het scoremodel van een punt

per item (ongeacht het discriminatievermogen) het voor het onderwijs het meest begrijpelijke en dus bruikbare model is.

Door middel van twee schalingen zijn de items van de toetsen IEP LVS Technisch Lezen 6b, 6a en 5b op dezelfde parameterschaal gezet. De toets 5b Technisch Lezen is een bestaande toets die in 2020 is verantwoord. Door 5b samen te schalen met de nieuwe toetsen 6a en 6b wordt een doorlopende lijn van vaardigheidsschattingen voor de vaardigheid Technisch Lezen op de verschillende niveaus gerealiseerd. In de eerste schaling zijn alleen de items van het onderdeel Leestechiek en in de tweede schaling zijn zowel de items Leestechiek als Leesexpressie van de zes IEP LVS toetsen Technisch Lezen gezamenlijk gekalibreerd.

In het document Itemparameters IEP LVS Technisch Lezen (bijlage 4) zijn het aantal afnames, de moeilijkheid (β), de meetfout van de moeilijkheid ($SE(\beta)$) en de iteminformatie (ook wel Fisher-informatie genoemd), bij de gemiddelde vaardigheid van de afnamegroepen van de IEP LVS toetsen Technisch Lezen 6b, 6a en 5b in een tabel weergegeven. Deze itemeigenschappen zijn gebaseerd op de gezamenlijke kalibratie van de items Leestechiek en Leesexpressie. Voor Leestechiek geldt dat alle items als of goed (score 1) of fout (score 0) beoordeeld kunnen worden waardoor er één parameter per item in de kalibratie wordt geschat. Bij de items van Leesexpressie is sprake van steeds twee parameterschattingen omdat twee mogelijke scorepunten boven nul voor deze items toegekend kunnen worden. Leesexpressie kent de beoordeelpunten onder niveau, op niveau en boven niveau. In bijlage 4 worden alle parameterschattingen gerapporteerd met de bijhorende scorepunten. De vaardigheidsschaal is genormeerd door van de afnamegroep van de 5b toets de gemiddelde vaardigheid en de standaarddeviatie hiervan op respectievelijk 0 en 1 te zetten.

In paragraaf 7.3 wordt een methode besproken om de nauwkeurigheid van de parameterschattingen te beoordelen (Evers et al., 2010). Deze methode bestaat eruit om de nauwkeurigheid van de parameterschattingen na te gaan aan de hand van de constante 'c', die de relatie weergeeft tussen de standaardfout van de moeilijkheidsparameter van een item ($SE(\beta)$) en de standaarddeviatie van de vaardigheidsverdeling van de kalibratiepopulatie (σ_{θ}). Volgens het COTAN-beoordelingssysteem (Evers et al., 2010) worden waarden van c lager of gelijk aan 0,2 als 'goed' beoordeeld en waarden tussen 0,3 en 0,4 als 'voldoende'. Zoals in paragraaf 6.2 is uitgelegd, de moeilijkheidsparameters van alle items Leestechiek en Leesexpressie van de IEP LVS toetsen Technisch Lezen 6b, 6a en 5b zijn ook gezamenlijk gekalibreerd. De nauwkeurigheid van de parameterschattingen is onderzocht door de c-waarden te berekenen voor de parameterschattingen die zijn voortgekomen uit deze gezamenlijke kalibratie. De gemiddelde waarde van de constante c, berekend over alle items in de kalibratie, is veel lager dan de vereiste waarde van 0,2. Met de uitzondering van één item is de nauwkeurigheid van de parameterschattingen voor alle items goed ($c < 0,2$). Eén item Leestechiek stijgt met 0,398 uit boven een c waarde van 0,2. De nauwkeurigheid van de parameterschatting van dit item is voldoende.

Om modelpassing te onderzoeken, mede in het kader van begripsvalidering, is er een Differential Item Functioning (DIF) analyse uitgevoerd in Lexter. De DIF van de items die in meer dan 1 LVS boekje voorkomen is berekend tussen de verschillende boekjes op basis van de kalibratie van de IEP LVS toetsen. Het betreft hier de 20 ankeritems tussen de opeenvolgende niveautoetsen. De resultaten zijn in bijlage 6 toegevoegd aan deze verantwoording. Ieder van deze ankeritems is in 2 boekjes opgenomen geweest ($df = 1$). In bijlage 7 is per item de Lagrange Multiplier statistiek (LM) weergegeven, het aantal

vrijheidsgraden (df), de overschrijdingskans (Prob), het absolute verschil (Abs.Dif) en is aangegeven in welke boekjes het item was opgenomen.

Omdat significantie altijd gevoelig is voor de grootte van de steekproef is het met name informatief om naar de absolute verschillen te kijken. Geen van de 20 ankeritems heeft een absoluut verschil van groter dan 0,10. Het gemiddelde absolute verschil van de 20 ankeritems is 0,02 (max=0,05; min=0,00). Op grond van deze resultaten kan er gesteld worden dat de observaties op de ankeritems tussen de IEP LVS toetsen Technisch Lezen n 6b, 6a en 5b weinig van de verwachting verschillen, ongeacht het boekje, wat een goede modelpassing aannemelijk maakt.

Voor de selectie van de ankeritems zijn dezelfde criteria gebruikt als voor de andere items (H6.1). Er is dus gekeken naar een gevarieerde inhoud van de ankeritems, de rit-waarde en de p-waarde. Er zijn bij de ankeritems enkele uitzonderingen gemaakt met betrekking tot de rit- en p-waarde. Met name bij de ankeritems naar het lagere niveau. Bij deze items is de p-waarde vaak hoog de rit-waarde daarmee wat laag. De ankeritems zijn ook opgenomen in Bijlage 5 (TIA's).

Conclusie:

Het onvolledige maar 'verbonden' dataverzamelsdesign is adequaat. Op aspect S2 wordt aan de toetsen IEP LVS Technisch Lezen n 6a en 6a het volgende oordeel toegekend: **'voldoende'**

S3 In het geval van een observatie-instrument: is er sprake van een adequate steekproef van observatoren en randvoorwaarden waaronder de observatie wordt uitgevoerd?

Bevindingen:

n.v.t.

Conclusie:

n.v.t.

S4 Er is een handleiding met duidelijke instructies voor de leerkracht over het zo objectief mogelijk uitvoeren en weergeven van de observaties door de leerkracht.

Bevindingen:

n.v.t.

Conclusie:

n.v.t.

Normering

N1.1 Is de standaardbepalingsmethode gemotiveerd en op de juiste wijze uitgevoerd?

Bevindingen:

De procedure die tijdens deze standaardsetting gevolgd is, is een combinatie van de Angoff methode en de Bookmark. In de Angoff methode wordt de cesuur bepaald door experts uit het betreffende vakgebied. Dit gebeurt in drie stappen:

1. De experts worden goed geïnformeerd. Zij ontwikkelen een beeld van de 'borderline candidate', ofwel de leerling die het niveau van de toets net behaald. Daarnaast worden de experts ook geïnformeerd over de consequenties van de niveau-uitspraak voor de leerling en de betekenis daarvan voor het onderwijs.
2. De experts geven individueel per item aan hoeveel procent van de 'borderline candidates' het item goed zullen maken.
3. Er wordt een discussie gevoerd tussen de experts om overeenstemming te bereiken over de resultaten van stap 2. Deze overeengekomen uitspraak resulteert vervolgens in een cesuur op de toets.

De tweede stap van de Angoff methode is een lastige stap voor onervaren standaardsetters. Ter ondersteuning van de standaardsetters met beperkte ervaring met het uitvoeren van een standaardsetting, is gekozen om die stap te vervangen door een deel van de Bookmark methode. Als tweede stap hebben de standaardsetters individueel in een lijst van items die op volgorde van makkelijk naar moeilijk lagen, een 'bookmark' geplaatst. De 'bookmark' betekent dat leerlingen die het niveau net beheersen de items tot de 'bookmark' goed moeten maken en de items na de 'bookmark' niet goed zouden hoeven te maken. Meer gedetailleerd doorlopen de standaardsetters de volgende stappen:

Stap 1: Informeren van standaardsetters

De standaardsetters zijn voorbereid door middel van een instructie. Zij kregen een toelichting over de aanleiding en het doel van de standaardsetting. Vervolgens hebben zij voor elke niveau (6a en 6b) de niveauomschrijvingen (bijlage 3) en de toetsmatrijs (bijlage 1, hoofdstuk 3.4) goed bekeken.

Stap 2a: 'Bookmark' plaatsen in onderdeel Leestechiek

Daarna hebben zij de 50 items (de woorden die door de leerkrachten tijdens het lezen van de tekst worden beoordeeld of deze correct gelezen zijn) van het onderdeel Leestechiek per toets op volgorde van makkelijk naar moeilijk gelegd. De standaardsetters kregen hierbij geen psychometrische informatie over de items en de toets. Op deze manier baseerden de standaardsetters hun oordeel puur op de inhoud van de items en werden zij aan het denken gezet over het niveau en de moeilijkheid van de items van de verschillende niveaus. Hierna moesten de standaardsetters bepalen hoeveel items een leerling die de doelen uit de toets n t beheerst goed moet lezen. Dit deden ze door een 'bookmark' te plaatsen in de door henzelf gemaakte volgorde van items van makkelijk naar moeilijk uit stap 1. Daarnaast hebben de standaardsetters aangegeven of zij akkoord waren met onderstaand principe voor het onderdeel Leestechiek. Alle standaardsetters waren akkoord met dit principe: Een leerling die het niveau beheerst, zou theoretisch gezien de woorden onder het niveau allemaal goed moeten lezen. Van de woorden boven het niveau hoeft de leerling er geen  en goed te lezen.

Stap 2b: 'Bookmark' plaatsen in onderdeel Leesexpressie

De standaardsetters hebben in stap 1 ook de beoordelingsmodellen voor het onderdeel Leesexpressie goed bekeken. In tegenstelling tot het onderdeel Leestechiek worden de aspecten niet gescoord als 'goed' of 'fout', maar beoordelen de leerkrachten de leerling op verschillende aspecten met een beoordeling 'onder niveau', 'op niveau' of 'boven niveau'. In deze stap hebben de standaardsetters aangegeven welke beoordeling een leerling minimaal moet krijgen om te kunnen zeggen dat de leerling de doelen uit de toets n t beheerst. Vervolgens hebben zij een uitspraak gedaan over of zij vinden dat een leerling mag compenseren tussen de verschillende aspecten die worden beoordeeld binnen het onderdeel Leesexpressie.

Stap 2c: Combineren van Leestechiek en Leesexpressie tot niveaubepaling Technisch Lezen

In de laatste stap in de voorbereiding hebben de standaardsetters aangegeven wat de relatie tussen de twee onderdelen zou moeten zijn om een goede niveaubepaling op de gehele vaardigheid Technisch Lezen te kunnen doen. Ze hebben aangegeven of ze vonden dat er   n onderdeel voorwaardelijk zou moeten zijn voor het andere onderdeel, of ze vonden dat leerlingen konden compenseren tussen de onderdelen en in welke verhouding beide onderdelen volgens hen zouden moeten meetellen.

Stap 3: Bereiken van consensus

De toetsen Technisch Lezen bestaan uit twee onderdelen (leestechiek en leesexpressie) waardoor de niveaubepaling een complex geheel is. Stap 1 t/m 2c van de vier individuele standaardsetters zijn daarom geanalyseerd en globaal uitgewerkt op papier door de procesbegeleider. Deze uitwerking is door de standaardsetters gelezen, waarna er een online discussie middels videobellen is geweest met de standaardsetcommissie en de procesbegeleider om dit te bespreken. Alle standaardsetters hebben na de discussie op papier ontvangen waarover consensus is bereikt en hebben daarop nogmaals hun akkoord gegeven.

Conclusie:

De standaardbepalingsmethode is gemotiveerd en op de juiste wijze uitgevoerd. Op aspect N1.1 wordt aan de toetsen IEP LVS Technisch Lezen 6a en 6b het volgende oordeel toegekend: '**voldoende**'.

N1.2 Zijn de beoordelaars/vakdeskundigen/experts naar behoren geselecteerd en getraind?

Bevindingen:

Om zowel kennis uit het vakgebied Technisch Lezen als ervaring uit de praktijk samen te brengen, is gekozen om de standaardsettingscommissie te laten bestaan uit twee inhoudelijke experts en twee leerkrachten. Vier standaardsetters per standaardsetting is als voldoende geacht, omdat de toetsen van het leerlingvolgsysteem geen high-stakes toetsen zijn. De vereisten aan de inhoudelijke experts is dat zij gespecialiseerd zijn in het vakgebied Technisch Lezen. Daarnaast zijn zij onderwijskundig of taalkundig adviseur in het primair onderwijs. Ervaring als leerkracht in het basisonderwijs is voor de inhoudelijke experts niet vereist, maar wel wenselijk. De leerkrachten moeten recentelijk ervaring hebben in het lesgeven aan leerjaar 6 en zijn bij voorkeur op dit moment ook werkzaam als leerkracht van leerjaar 6. Bureau ICE heeft de rol van procesbegeleider in deze

standaardsetting. De standaardsettingscommissie is middels een werving binnen het eigen netwerk van Bureau ICE en sociale media geworven. Er is een screening uitgevoerd op het competentieprofiel van de geïnteresseerden. Dit heeft geresulteerd in een standaardsettingscommissie die voldoet aan bovenstaande vereisten. Alle standaardsetters hebben vooraf een geheimhoudingsverklaring ondertekend waarin zij verklaren dat zij de toetsinhoud vertrouwelijk behandelen.

In stap 1 van de gevolgde standaardsettingsprocedure werden de standaardsetters goed geïnformeerd over de 'borderline candidate' en ontwikkelden zij hierover een beeld. Ook werden zij in stap 1 geïnformeerd over de consequenties van de niveau-uitspraak voor de leerling en de betekenis daarvan voor het onderwijs

Conclusie:

De beoordelaars/vakdeskundigen/experts zijn naar behoren geselecteerd en getraind. Op aspect N1.2 wordt aan de toetsen IEP LVS Technisch Lezen 6a en 6b het volgende oordeel toegekend: **'voldoende'**.

N1.3 Is er voldoende overeenstemming tussen de beoordelaars?

Bevindingen:

In deze discussie zijn de onderwerpen 'Afhankelijkheid onderdelen Leestechneik en Leesexpressie', 'cesuur Leestechneik' en 'cesuur Leesexpressie' aan bod gekomen. In paragraaf 7.1 wordt voor al deze drie onderdelen beschreven waarover overeenstemming is bereikt.

Conclusie:

Er is voldoende overeenstemming tussen de beoordelaars. Op aspect N1.3 wordt aan de toetsen IEP LVS Technisch Lezen 6a en 6b het volgende oordeel toegekend: **'voldoende'**.

N2.1 Zijn de normgroepen groot genoeg?

Bevindingen:

De toetsen zijn genormeerd voor de afnamemomenten a (eerste helft van het schooljaar) en b (tweede helft van het schooljaar) in leerjaar 6. De grootte van de normeringsgroepen kan afgelezen worden uit Tabel 5.1 en is van voldoende grootte.

Conclusie:

De normgroepen zijn groot genoeg.

Op aspect N2.1 wordt aan de toetsen IEP LVS Technisch Lezen 6a en 6b het oordeel **'voldoende'** toegekend.

N2.2 Zijn de normgroepen representatief?

Bevindingen:

De representativiteit van de steekproeven werd hierboven onder aspect S1 besproken en daar werd geconstateerd dat de representativiteit van de normeringspopulatie op de achtergrondvariabelen niet optimaal was om als normpopulatie voor het gehele reguliere

basisonderwijs (doelpopulatie) te fungeren. Volgens de auteurs was het echter ook niet per se noodzakelijk dat de normeringspopulatie ook een normpopulatie is en zijn de representativiteitseisen aan een normpopulatie in het onderhavige onderzoek dan ook niet van toepassing. De argumentatie hiervoor is dat in dit normeringsonderzoek absolute cesuren op de toetsen werden vastgesteld, waarvoor de samenstelling van de normeringspopulatie volgens de auteurs van ondergeschikt belang is. Wel is het volgens de auteurs voor de bepaling van de kwaliteit van de items (hoofdstuk 6) en de standaardsetting (hoofdstuk 7) van belang vast te stellen dat de normeringspopulatie geen specifiek selecte groep is van de leerlingen in leerjaar 6 van het reguliere basisonderwijs.

Het is volgens de auteurs voor de bepaling van de gemiddelde groeifactor (zie hoofdstuk 9) echter wel van belang vast te stellen dat de normeringspopulatie een representatieve steekproef is, omdat daar een relatieve norm wordt bepaald. Voor het bepalen van de gemiddelde groeifactor kon volgens de auteurs representativiteit gegarandeerd worden door de data te filteren op twee criteria, namelijk ten eerste door records van SBO-scholen en van scholen van Bonaire uit het databestand te verwijderen en ten tweede door records waarvoor het toetsresultaat niet in een ontwikkelscore uitgedrukt wordt niet te gebruiken.

Conclusie:

De normgroepen zijn representatief. Op aspect N2.2 wordt aan de toetsen IEP LVS Lezen 6a en 6b het oordeel '**voldoende**' toegekend.

N2.3 Zijn de normen correct bepaald?

Bevindingen:

De gevolgde standaardsettingsprocedure past bij de data van de normeringspopulatie en er zijn ook schattingsfouten van de cesuren en percentages misclassificaties berekend.

Conclusie:

De normen zijn correct bepaald. Op aspect N2.3 wordt aan de toetsen IEP LVS Lezen 6a en 6b het oordeel '**voldoende**' toegekend

Betrouwbaarheid

B1 Zijn of worden de betrouwbaarheidsgegevens correct berekend?

Bevindingen:

In Bijlage 5 (TIA's 6a en 6b) worden de globale betrouwbaarheden (Cronbach's alpha) weergegeven voor de twee IEP LVS-toetsen Technisch Lezen (6a en 6b), welke zijn berekend met het programma TiaPlus. Deze coëfficiënt wordt in de psychometrische literatuur beschreven en als correct aangemerkt. Onder gebruikmaking van het programma Lexter worden in tabel 7.1a ook de lokale betrouwbaarheden, gemeten bij de cesuurpunten op de latente vaardigheidsschaal, weergegeven voor de toetsen 6a en 6b.

In bijlage 8 ('Algemene toelichting methode') wordt gedetailleerde uitleg gegeven hoe deze (conditionele) lokale betrouwbaarheden, gegeven een vaardigheidsniveau θ , kunnen worden berekend. Deze (conditionele) lokale betrouwbaarheid vertoont qua interpretatie grote overeenkomst met de globale betrouwbaarheidscoëfficiënt Cronbach's alpha uit de

klassieke testtheorie (KTT) en wordt in de psychometrische literatuur beschreven (Raju, Price, Oshima & Nering, 2007) en als correct aangemerkt.

Daarnaast worden in tabel 7.1a ook nog het percentage leerlingen berekend dat het betreffende niveau ten onrechte wel of niet heeft gehaald (classificatiefouten). In bijlage 5 ('Algemene toelichting methode') wordt op een correcte manier beschreven hoe deze classificatiefouten worden berekend. Omdat voor de berekeningen gebruik is gemaakt van bekende en algemeen beschikbare software (i.e., TiaPlus en Lexter), kunnen we ervan uitgaan dat de betrouwbaarheidsgegevens correct zijn berekend.

Conclusie:

De betrouwbaarheidsgegevens worden correct berekend. Op aspect B1 wordt aan de toetsen IEP LVS Technisch Lezen 6a en 6b het oordeel '**voldoende**' toegekend.

B2 Zijn de betrouwbaarheidsgegevens voldoende gezien de conclusies en eventuele beslissingen die met het instrument genomen worden?

Bevindingen:

In bijlage 4 (TIA's 6a en 6b) kan afgelezen worden dat de globale betrouwbaarheden in termen van Cronbach's alpha (interne consistentie betrouwbaarheden) voor de twee IEP LVS toetsen Lezen 6a en 6b gelijk zijn aan respectievelijk 0.81 en 0.79.

Het gemiddelde van de geschatte lokale betrouwbaarheid van de scorepunten binnen het bereik van voldoende beheersing is gelijk of hoger dan 0,819 van de drie IEP LVS toetsen Technisch Lezen. Het minimum van de lokale betrouwbaarheid van het bereik van scorepunten bij voldoende beheersing is 6a en 6b is gelijk aan of hoger dan 0,755. Aan het eis van het beoordelingssysteem van COTAN dat de minimale betrouwbaarheidscoëfficiënt van toetsen voor minder belangrijke beslissingen tenminste 0,70 moet zijn is dus bij de IEP LVS toetsen Technisch Lezen ruimschoots voldaan.

Verder laat tabel 7.2 nog zien dat de classificatiefouten (i.e., voor de vier IEP LVS-toetsen Lezen het percentage leerlingen dat het betreffende niveau ten onrechte wel of niet heeft gehaald) loopt van 6% tot 10% (hoe hoger de lokale betrouwbaarheid, hoe lager de classificatiefout). Deze percentages hebben betrekking op scores dicht bij een cesuur en er geldt dan ook dat het percentage misclassificaties bij een score verder van de cesuur af per definitie lager is. Omdat de berekende classificatiefouten in de context van de IEP LVS toetsen geen summatieve toetsen betreft waarop een leerling kan zakken of slagen, heeft een misclassificatie daarmee voor de leerling geen directe grote gevolgen. In combinatie met het feit dat de IEP LVS toetsen volgtoetsen zijn waar geen belangrijke beslissingen mee worden genomen, kan er geconcludeerd worden dat de classificatiefouten als acceptabel gezien kunnen worden.

Conclusie:

De betrouwbaarheidsgegevens zijn voldoende gezien de conclusies en eventuele beslissingen die met het instrument genomen worden. Op aspect B2 wordt aan de toetsen IEP LVS Technisch Lezen 6a en 6b het volgende oordeel toegekend: '**voldoende**'.

Validiteit

V1 Inhoudsvaliditeit: Dragen de items in het instrument bij aan de validiteit van het instrument (hierbij gaat het om aspecten als relevantie, objectiviteit en efficiëntie van de items)?

Bevindingen:

De toetsen Technisch lezen van het IEP LVS bestaan uit de toetsen Leestechneik en Leesexpressie. Iedere toets Leestechneik bestaat uit 40 items waarvan 5 items op het bovenliggende en 5 op het onderliggende niveau. Voor een vlotte en accurate woordherkenning (Referentiekader kenmerken van de taakuitvoering: techniek) wordt de decodeervaardigheid met de toetsen leestechneik getoetst om specifieke woorden op een bepaald niveau hardop te kunnen lezen. Via het lezen van teksten waarin de te toetsen woorden zijn opgenomen, worden de leerdoelen gemeten.

In de toetswijzer Technisch Lezen staat beschreven wat de leerdoelen zijn en het beheersingsniveau.

Leestechneik kent de volgende categorieën:

- Klankzuivere woorden
- Niet-klankzuivere woorden
- Vaste lettercombinaties
- Samenstellingen en afleidingen
- Leenwoorden

De leerdoelen die getoetst worden per niveau, bepalen welke categorieën en hoeveel woorden van iedere categorie worden aangeboden. Via de toetsmatrijzen is een duidelijke opbouw zichtbaar van de toetsen 3A t/m 6B.

Doelen gericht op de algehele leesprestatie worden gemeten via de toetsen Leesexpressie. Beoordeling vindt plaats op:

- Vloeiendheid
- Intonatie leestekens
- Zinsmelodie
- Directe rede
- Fouten en correctie
- Leestempo

De Matrijs Technisch lezen – leesexpressie geeft per toets weer of een aspect voorwaardelijk, aanvullend of ter indicatie is. Voor leerjaar 6 is 'directe rede' als zesde aspect toegevoegd. Bij niveau 6 worden alle aspecten beoordeeld.

Om het onderdeel leesexpressie te kunnen halen, zijn de aspecten vloeiendheid, intonatie van leestekens, directe rede en fouten & correctie voorwaardelijk. Het aspect zinsmelodie is aanvullend staat in de Toetswijzer op pagina 11.

Dit klopt niet met wat er in de tabellen 18 en 20 staan. Voor 6a wordt aangegeven in de tabel dat alleen het aspect directe rede aanvullend is en in tabel 6b dat alle aspecten voorwaardelijk zijn.

Leestempo wordt in alle toetsen ter indicatie opgenomen zoals ook in de tabellen wordt aangegeven en weegt niet mee in de beoordeling. De tijdsmarge kan wel aangegeven worden.

Bij de toetsen Leestechneik en bij de toetsen Leesexpressie 6a en 6b wordt een kwantitatieve duiding gegeven bij 3 van de 6 aspecten.

Bijvoorbeeld bij intonatie van leestekens:

- Minder dan 60% van de leestekens zijn herkenbaar door middel van intonatie.
- 60%-80% van de leestekens zijn herkenbaar door middel van intonatie
- Meer dan 80% van de leestekens zijn herkenbaar door middel van intonatie.

Bij de beoordeling van leesexpressie bij het aspect directe rede wordt aangegeven dat een andere eis wordt gesteld bij de beoordeling per niveau-beschrijving:

- De aanhalingstekens en tekstuele aanwijzingen worden niet verwerkt in de uitingen van de leerling (onder niveau).
- De aanhalingstekens worden verwerkt in de uitingen van de leerling, maar de tekstuele aanwijzingen niet (op niveau).
- De aanhalingstekens en tekstuele aanwijzingen worden beide verwerkt in de uitingen van de leerling (boven niveau).

Het lijkt dat de leraar hier zelf een maat moet bepalen. Wat als aanhalingstekens soms wel en soms niet worden verwerkt in de uitingen? Is het dan op of onder niveau?

Audiofragmenten bij alle toetsen zijn een goede ondersteuning voor de leraar om een meer objectieve beoordeling te kunnen geven.

Voor technisch lezen, wordt anders dan voor begrijpend lezen, geen eisen gesteld aan het soort teksten. Het is een goede keuze dat er gekozen is voor fictionele en informatieve teksten omdat dit de teksten zijn die kinderen bij het lezen op school en thuis ook het meest tegenkomen.

Er is bij de uitwerking van de toetsen leestechneik niet gekozen voor het lezen van woordrijtjes. De leerlingen krijgen zinnen aangeboden. De teksten 6a en 6b van leestechneik bestaan zoveel mogelijk uit logisch lopende zinnen die een informatieve/fictieve tekst vormen. Doordat bepaalde categorieën op een bepaald niveau in de tekst moeten voorkomen, zit er wel verschil in aantrekkelijkheid en toegankelijkheid van teksten in vergelijking met teksten die voorkomen in de IEP toetsen van (begrijpend) lezen. De teksten voor de toetsen Leestechneik zijn wat meer kunstmatig uitgewerkt. Het lezen van zinnen heeft wel de voorkeur boven het lezen van losse woordrijtjes dus daar is een goede keuze in gemaakt.

Totaalopmerking over de uitwerking van de toetsen Technisch lezen:

Voor het leesbegrip is behalve tempo en correctheid ook expressie bepalend. Veelal is dit niet terug te vinden in de monitoring van het leesonderwijs. Leerlingen krijgen veelal feedback op snelheid en correctheid. Het is een meerwaarde dat de IEP toetsen van Technisch lezen zowel uitgewerkt zijn op leestechneik als op leesexpressie.

Conclusie:

'Voldoende'

V2 Constructvaliditeit: Meet het instrument in zijn geheel datgene wat het beoogt te meten?

Bevindingen:

Voor technisch lezen zijn er t/m groep 6 leerdoelen uitgewerkt gebaseerd op de aangeboden leesmoelijkheden. Woorden die aangeboden worden, corresponderen met de leesmoelijkheden die een leerling zou moeten beheersen volgens het toetskader Technisch Lezen (maar zie ook Leerstoflijnen beschreven, SLO 2010). Na groep 6 gaat het om het onderhouden van de leesvaardigheden daarom worden er na groep 6 geen toetsen Technisch lezen meer aangeboden.

Een toets Technisch lezen bestaat uit twee teksten. De eerste tekst is gericht op leesteknik en de tweede op leesexpressie. De leerling leest de teksten hardop voor en de afname van de toets vindt één op één plaats. In de digitale beoordeling van de toets is de afhankelijkheid tussen de beide onderdelen meegenomen. Als er geen zinvolle uit-spraak kan worden gedaan over het beheerste niveau op het onderdeel leesteknik omdat een leerling te laag scoort, dan wordt de toets gestopt en wordt het onderdeel leesexpressie niet meer afgenomen. De leraar krijgt dan als advies een toets van een lager niveau af te nemen. De onderbouwing die wordt aangegeven is inderdaad juist dat de decodeervaardigheid voorwaardelijk is voor de andere aspecten van de technische leesvaardigheid.

De woorden die bij de toets leesteknik worden gescoord, worden aangeboden in een tekst. Alleen de te beoordelen zogenaamde scorewoorden passend bij de gestelde leerdoelen worden gescoord. De scorewoorden zijn goed te herleiden naar de leerdoelen die leerlingen zouden moeten beheersen op het betreffende te toetsen niveau.

In overleg met de inhoudelijke experts van ICE is ervoor gekozen om leesteknik zwaarder te laten meewegen dan leesexpressie omdat dit de basis is van de technische leesvaardigheid. Gezien de insteek van de toetsen kan dit onderschreven worden.

In paragraaf 7.3 ('Passing van het meetmodel en nauwkeurigheid van de parameterschattingen') was al aannemelijk gemaakt dat er sprake was van een passing van het meetmodel en er dus mag worden uitgegaan van unidimensionaliteit, hetgeen impliceert dat aan de noodzakelijke (maar niet voldoende) voorwaarde van constructvaliditeit wordt voldaan. In hoofdstuk 8 wordt aanvullend onderzoek verricht naar argumenten die pleiten voor de constructvaliditeit van de IEP LVS-toetsen Lezen.

Om de constructvaliditeit van de IEP LVS toetsen Technisch Lezen te evalueren zijn de volgende (inter)correlationele onderzoeken zijn uitgevoerd:

- correlaties tussen de onderdelen Leesteknik en Leesexpressie binnen de toetsen
- correlatie tussen a- en b-toetsen per onderdeel
- correlaties tussen toetsen van verschillende vaardigheden

Uit tabel 8.1 valt af te lezen dat er een middelmatig verband is tussen de onderdelen Leesteknik en Leesexpressie. Hoewel om constructvaliditeit aannemelijk te maken statistisch hoge tot zeer hoge verbanden prijzenswaardig zijn, zijn de 'slechts' middelmatige verbanden geenszins onverwacht. Inhoudelijk zijn beide onderdelen weliswaar een essentieel aspect van Technisch Lezen, maar beide onderdelen verschillen qua karakteristiek aanzienlijk. Tevens verschilt de beoordelingsystematiek van beide onderdelen.

In tabel 8.2 zijn de correlatiecoëfficiënten weergegeven tussen Leestechniek in het eerste deel van het schooljaar en Leestechniek in het tweede deel van het schooljaar. Dit is ook gedaan voor Leesexpressie.

De correlatie tussen gelijke onderdelen van de toetsen Technisch Lezen van twee opeenvolgende toetsen Technisch Lezen zijn middelmatig. Deze correlaties zijn structureel hoger dan met de twee vaardigheden Lezen en Rekenen binnen hetzelfde leerjaar (tabel 8.3). Dit ondersteunt de aannemelijkheid van de constructvaliditeit.

De correlaties van Leesexpressie (of de samengestelde score van Leestechniek en Leesexpressie) en de vaardigheden van Lezen en Rekenen zijn toegevoegd aan tabel 8.3 op pagina 26.

Conclusie:

De gerapporteerde resultaten in Hoofdstuk 8 (Constructvaliditeit) van de WV vormen een psychometrische ondersteuning voor de constructvaliditeit van de toetsen IEP LVS Lezen 6a en 6b en het is aannemelijk dat er dus wordt gemeten wat men beoogt te meten, namelijk technisch lezen bij leerjaar 6. Op aspect V2 wordt aan de toetsen IEP LVS Technisch Lezen 6a en 6b het volgende oordeel toegekend: **'voldoende'**.

Het volg-aspect

Va1 Is er een voldoende empirische onderbouwing van de schaal waarop de groei van een leerling wordt uitgedrukt? Wordt groei op een correcte manier gemeten?

Bevindingen:

Om de ontwikkeling van een leerling zinvol te kunnen volgen moeten alle toetsen van één vaardigheid die een leerling in zijn/haar schoolloopbaan maakt op één en dezelfde schaal een uitslag geven. Dit wordt in het IEP LVS leerjaar 3 tot 6 gerealiseerd doordat alle toetsen een uitspraak doen over de beheersing van inhoudelijk doelen. Deze doelen zijn door het onderwijs per vaardigheid geformuleerd en kennen een hiërarchische opbouw.

In tabel 9.1 wordt grafisch inzichtelijk gemaakt hoe een bepaalde toetsscore, door gebruik te maken van drie vaste punten (bodem, cesuur en plafond) omgezet kan worden in een ontwikkelscore. De gebieden tussen de bodem en de cesuur, en tussen de cesuur en het plafond, zijn vervolgens lineair verdeeld bij de berekening van de ontwikkelscorepunten met toepassing van de in de standaardsetting bepaalde weging van 60%-40% van de onderdelen Leestechniek en Leesexpressie (hoofdstuk 7.1). De berekende ontwikkelscore wordt altijd afgerond op een heel getal weergegeven.

Conclusie:

Er is voldoende empirische onderbouwing van de schaal waarop de groei van een leerling wordt uitgedrukt. Op aspect Va1 wordt aan de toetsen IEP LVS Technisch Lezen 6a en 6b het volgende oordeel toegekend: **'voldoende'**.

Va2 Wordt de betrouwbaarheid van de groei op die schaal correct weergegeven?

Bevindingen:

De (conditionele) lokale betrouwbaarheid voor ieder scorepunt wordt op dezelfde manier geschat als de (conditionele) lokale betrouwbaarheid van de cesuurpunten (op de thetaschaal) voor de twee IEP LVS-toetsen en is beschreven in de tabellen 7.1b en 7.1c (zie bijlage 8 voor uitleg over de gedetailleerde berekening). Deze (conditionele) lokale betrouwbaarheid vertoont qua interpretatie grote overeenkomst met de globale betrouwbaarheidscoëfficiënt Cronbach's alpha uit de klassieke testtheorie (KTT) en wordt in de psychometrische literatuur beschreven en als correct aangemerkt.

Conclusie:

De betrouwbaarheid van de groei op de ontwikkelscoreschaal wordt correct weergegeven. Op aspect Va2 wordt aan de toetsen IEP LVS Technisch Lezen 6a en 6b het oordeel '**voldoende**' toegekend.

Va3 Worden er voldoende gegevens verstrekt (aan de gebruiker) over hoe groei geïnterpreteerd dient te worden?

Bevindingen:

De interpretatie van de behaalde resultaten op de toetsen IEP LVS Lezen 6a en 6b wordt gevisualiseerd door de behaalde ontwikkelscores grafisch weer te geven op twee manieren. Ten eerste geeft de voortgangsgrafiek (zie figuur 9.2) de vaardigheidsontwikkeling weer in de tijd door de toetsresultaten van alle toetsen, die in de leerjaren 3, 4, 5 en 6 zijn afgenomen, op één en dezelfde schaal uit te drukken. Ten tweede wordt in de leergroemeter (zie figuur 9.4) de leergroei van een leerling afgezet tegen de gemiddelde groeifactor van de normeringspopulatie en geeft dus aan hoe snel een leerling groeit ten opzichte van de normeringspopulatie. Het verschil in toetsresultaat, uitgedrukt in ontwikkelscorepunten, tussen twee of meer opeenvolgende toetsmomenten duidt hierbij de leergroei van een leerling aan. De gemiddelde groeifactor wordt berekend om een relatieve beoordeling van de leergroei per leerling te visualiseren. In paragraaf 9.2 ('Leergroei') wordt gedetailleerd beschreven hoe de gemiddelde groeifactor van de normeringspopulatie wordt berekend.

Om de absolute leergroei te duiden is de representatie van de referentiepulatie van ondergeschikt belang, maar voor de vergelijking van de leergroei met de gemiddelde groeifactor (weergegeven in de leergroemeter, zie figuur 9.4) is de representativiteit wel van belang. Dit betekent dat voor de toekomst de gemiddelde groeifactor jaarlijks geëvalueerd gaat worden aan de hand van representatieve afnamedata en indien nodig wordt aangepast.

De interpretatie van de leergroei van leerlingen wordt voor de leerkrachten, leerlingen en ouders/verzorgers ondersteund door de twee grafische weergaven, de voortgangsgrafiek en de leergroemeter, gecombineerd te gebruiken in de leergroemeter (zie fig. 9.4). Voor leerkrachten zijn de leervorderingen van een leerling ook digitaal beschikbaar in het IEP LVS, welke door de leerkrachten ook geprint kunnen worden in de vorm van de IEP LVS Talentenkaart (zie bijlage 10 'Talentenkaart' voor een voorbeeld van een Talentenkaart). Daarnaast hebben leerkrachten ook per toetsresultaat inzicht in de scores (percentage goed beantwoorde items) op de verschillende inhoudelijke domeinen en kunnen de

gegeven antwoorden van de leerling inzien, hetgeen het formatief gebruik en de bruikbaarheid van de IEP LVS toetsen bevordert.

Conclusie:

Er worden voldoende gegevens verstrekt (aan de gebruiker) over hoe groei geïnterpreteerd dient te worden. Op aspect V31 wordt aan de toetsen IEP LVS Technisch Lezen 6a en 6b het volgende oordeel toegekend: **'voldoende'**.

Inzicht in leervorderingen

I1 Levert de aanbieder een geschreven toelichting bij de leervorderingen van de leerling die (ook) voor ouders /verzorgers/voogden/docenten begrijpelijk is?

Bevindingen:

Er is een digitale handleiding van het IEP LVS. Daarin wordt de leraar via een achttal stappen wegwijs gemaakt in het werken met het LVS. Daarin wordt ook in stap 6 inzichtelijk via voorbeelden aangegeven welke mogelijkheden het systeem biedt om leergroei van individuele leerlingen in kaart te brengen. Ook wordt inzicht gegeven hoe de resultaten van de groep door groepsvoortgangsgrafieken afgezet kan worden tegen het landelijk gemiddelde. Resultaten kunnen gegroepeerd worden op bijvoorbeeld ontwikkelscore, domein en niveau.

In het leerlingenprofiel kun je naast de leergroei-meter en individuele voortgangsgrafiek van het betreffende schooljaar ook nog de resultaten van voorgaande schooljaren bekijken, ook per vaardigheid. Dit biedt een bredere kijk op hoe de ontwikkeling van de betreffende leerling verlopen is.

In de toetsen 6A en 6B worden de ontwikkelscores afgezet tegen de referentieniveaus.

Naast de handleiding en handreiking kan ook het document Toelichtingen Technisch Lezen groep 3 t/m 6 erbij worden gebruikt door de leraar waar inzicht wordt gegeven in het toetskader. Het geeft de leraar een duidelijk beeld wat er qua niveau getoetst wordt, dus de inhoud van de toets, geïllustreerd met voorbeeldopgaven van de vaardigheden per domein.

De Talentenkaart is een overzichtelijke uitwerking om voor individuele kinderen en hun ouders/verzorgers inzicht te bieden in de voortgang van hun kind. Niet alleen op taal en rekenen maar ook op sociaal-emotioneel gebied. Leraren kunnen zelf een aantal persoonlijke opmerkingen opnemen over bijvoorbeeld de leerling zijn leerhouding en leerresultaten. Daarnaast wordt inzicht gegeven in de toetsresultaten (score), een voortgangsgrafiek en een grafische weergave van de leergroei (leergroei-meter) en detailscores op o.a. technisch lezen. Dit biedt inzicht in waar een kind op dat moment staat en of er sprake is van leergroei.

Voor leraren is een leeswijzer bij de Talentenkaart uitgewerkt voor de groepen 6 t/m 8 waar ook al wordt vooruitgekeken op het schooladvies voor de kinderen in de groepen 7 en 8. Daarnaast is duidelijk en in het kort uitgewerkt wat referentieniveaus zijn en hoe de niveaus afgezet zijn tegen de IEP ontwikkelscores.

Er was ter beoordeling geen leeswijzer van de Talentenkaart toegevoegd voor ouders maar we gaan ervanuit, omdat deze er ook is voor ouders voor de leerjaren 3 t/m 5 om leervorderingen te interpreteren, die er ook is voor de ouders/verzorgers van leerjaar 6. De leeswijzer vraagt uitleg van de leraar voor ouders die niet gewend zijn om grafieken af te lezen. En ook om te weten welke doelen de leraar voor ogen heeft met hun kind en hoe dit zich verhoudt tot de getoonde resultaten.

Op de site van bureau ICE is een aparte pagina voor ouders met compacte informatie over o.a. het LVS.

Conclusie:

De aanbieder (i.e., Bureau ICE) levert een geschreven toelichting bij de leervorderingen van de leerling die voor docenten begrijpelijk is. Op aspect I1 wordt aan de toetsen IEP LVS Technisch Lezen 6a en 6b het oordeel '**voldoende**' toegekend

I2 Is er een evaluatie van de leervorderingen en worden op basis van deze evaluatie vervolgstappen geformuleerd?

Bevindingen:

In de handleiding wordt voor het interpreteren van toetsresultaten via een link verwezen naar de betreffende handreiking hierover. Daar is in een tabel een schematische aanpak uitgewerkt hoe je na kunt gaan of je als leraar op basis van de behaalde toetsresultaten en ontwikkelscore door de leerling een passende toets hebt klaargezet en wat je vervolgens kunt beslissen.

In de digitale handleiding van het IEP LVS wordt bij stap 6 uitgelegd hoe resultaten gegroepeerd kunnen worden op o.a. resultaat, niveau en domein. Dit geeft op verschillende manieren inzicht in hoe de leergroei van de individuele leerling en de groep is geweest, maar biedt ook inzicht welke doelen, vertaald in een aanpak, verder gestimuleerd en opgepakt kunnen worden door de leraar voor de groep en op individueel niveau. De resultaten per individu en groep kunnen in een grafiek afgezet worden tegen het landelijk gemiddelde of de schoolweging. Het is ook mogelijk in het kader van kwaliteitszorg dwarsdoorsnedes te maken op leerjaar- en groepsniveau om de opbrengsten op schoolniveau te evalueren.

Bij de toetsen leesexpressie wordt leestempo in alle toetsen ter indicatie opgenomen en telt niet mee bij de beoordeling. Omdat leestijd wel feedback kan geven over de leesvaardigheid van de leerling, is er bij dit aspect aangegeven tussen welke leestijden de leerling heeft gelezen.

Bij het voortgezet lezen zijn leestempo en automatisering wel belangrijk. In de toetsen leesexpressie bepalen vloeiendheid en de hoeveelheid fouten die de leerling maakt bij de tekst mede het leestempo. Een lage score zal er ook toe leiden dat een leerling langer dan gemiddeld met een tekst bezig is. Dit wordt ook duidelijk aangegeven en is belangrijk om te weten bij de te kiezen aanpak op basis van de toetsresultaten.

Het document Toelichtingen Technisch Lezen groep 3 t/m 6 kan de leraar, omdat het een duidelijk beeld geeft wat er qua niveau wordt getoetst, ondersteunen bij de keuze voor vervolgstappen t.a.v. aspecten van lezen.

Er wordt nadrukkelijk aangegeven dat naast het vergelijken met een norm het aangeboden aanbod en de gestelde ambities/doelen van belang zijn om de getoonde leergroei te 'wegen' en dus ook hier consequenties uit te trekken voor de vervolgstappen. Het is goed dat op verschillende plekken in de handreiking en in de andere documenten de cruciale rol van de leraar wordt benadrukt om te komen tot een juiste interpretatie en vervolgaanpak.

Voor bestuurders is er een Bovenschoolse IEP module Analyse en Kwaliteitszorg om trendanalyses, dwarsdoorsnedes en eindresultaten te bekijken van de scholen die onder de stichting vallen en met het IEP LVS werken. Op stichtingsniveau is het op deze manier ook mogelijk om te evalueren en gezamenlijk na te gaan welke aanpakken in de verschillende schoolcontexten tot goede resultaten leiden en op welke wijze schoolteams ook van elkaar kunnen leren.

De leerkracht wordt ondersteund bij de interpretatie van de vaardigheidsontwikkeling van de leerling door het gecombineerd gebruik van de twee grafische weergaven, de voortgangsgrafiek (zie fig. 9.2) en de leergroeimeter (zie fig. 9.4). De leerkracht kan hiermee evalueren in welke mate de leerling ten opzichte van zijn/haar verwachting en/of ten opzichte van de verwachte groeifactor zich ontwikkelt en kan hij/zij inschatten hoe waarschijnlijk het is dat de leerling het beoogde streefniveau zal gaan bereiken. In de 'Handreiking interpreteren toetsresultaten' (bijlage 9) worden leerkrachten geholpen bij de interpretatie van de ontwikkelscores en krijgen zij advies over het bepalen of een toets 'passend' was qua niveau voor de leerling.

Conclusie:

Er is een evaluatie van de leervorderingen en op basis van deze evaluatie worden vervolgstappen geformuleerd. Op aspect I2 wordt aan de toetsen IEP LVS Technisch Lezen 6a en 6b het oordeel '**voldoende**' toegekend.

Referentieniveaus

R1 Sluit de inhoud van de toets aan op de kennis en vaardigheden zoals omschreven in de referentieniveaus van het betreffende domein (voor toetsen vanaf groep 6)?

Bevindingen:

In hoofdstuk 7 wordt aangegeven waar de cesuren liggen voor de niveaus A1 en A2 van het CEFR. In de Toetswijzer (bijlage 1) wordt expliciet aangegeven dat de IEP LVS toetsen voor de leerjaren 6, 7 en 8 (met uitzondering van Technisch lezen leerjaar 6) zijn gebaseerd op het Referentiekader taal en rekenen (Meijerink et al., 2009). Er wordt dus een uitzondering gemaakt voor Technisch Lezen.

Conclusie:

Op aspect R1 wordt aan de toetsen IEP LVS Technisch Lezen 6a en 6b het oordeel '**voldoende**' toegekend.

3. Verzamelstaat

Kwaliteitsaspect	Code	Oordeel
De kwaliteit van de steekproef	<i>S1</i>	Voldoende
	<i>S2</i>	Voldoende
	<i>S3</i>	n.v.t.
	<i>S4</i>	n.v.t.
Normering	<i>N1.1</i>	Voldoende
	<i>N1.2</i>	Voldoende
	<i>N1.3</i>	Voldoende
	<i>N2.1</i>	Voldoende
	<i>N2.2</i>	Voldoende
	<i>N2.3</i>	Voldoende
Betrouwbaarheid	<i>B1</i>	Voldoende
	<i>B2</i>	Voldoende
Validiteit	<i>V1</i>	Voldoende
	<i>V2</i>	voldoende
Volg-aspect	<i>Va1</i>	Voldoende
	<i>Va2</i>	Voldoende
	<i>Va3</i>	Voldoende
Inzicht in leervorderingen	<i>I1</i>	voldoende
	<i>I2</i>	Voldoende
Referentieniveaus	<i>R1</i>	voldoende

4. Literatuurlijst

- Bezdan, E., Binsbergen, M., Haitjema, T, Helsloot, J. & Laan, J. (2021). Verantwoording IEP LVS-toetsen Lezen. Culemborg: Bureau ICE.
- Evers, A., Lucassen, W., Meijer, R., & Sijtsma, K. (2010). COTAN Beoordelingssysteem voor de kwaliteit van tests. Amsterdam: Nederlands Instituut van Psychologen.