

1. Uitgangspunten van de toetsconstructie

Bij onderstaande beoordeling van de kwaliteitsaspecten met bijbehorende codes van het voornoemde beoordelingskader worden passages uit de Handleiding en Verantwoording veelal letterlijk vermeld. De Handleiding en Verantwoording bespreken de uitgangspunten van de toets, waaronder de afname, scoring, normering, interpretatie, rapportage en foutenanalyse en de psychometrische kenmerken van de toets..

Algemeen

De Schoolvaardigheidstoets (SVT) Spelling is onderdeel van het leerlingvolgsysteem van Boom test uitgevers. Het leerlingvolgsysteem van Boom test uitgevers Onderwijs bestaat uit de Schoolvaardigheidstoetsen Technisch Lezen, Begrijpend Lezen, Hoofdrekenen, Rekenen-Wiskunde en Spelling. Met de schoolvaardigheidstoetsen worden de leerlingen vanaf groep 3 tot en met 8 op de belangrijkste vaardigheden gevolgd.

De SVT Spelling bestaat uit twee verschillende onderdelen, te weten SVT Spelling Woorden en SVT Spelling Werkwoorden.

Meetpretentie

De SVT Spelling meet de spellingsvaardigheid van leerlingen zoals deze wordt aangeleerd in het Nederlandse basisonderwijs. Daarbij gaat het niet om een toetsing van de expliciete kennis van de spellingsregels, 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.

Doelgroep

De SVT Spelling Woorden is bestemd voor leerlingen in groep 3 tot en met 8 van het basisonderwijs. De SVT Spelling Werkwoorden is bestemd voor leerlingen in groep 6 tot en met 8 in het basisonderwijs.

Gebruiksdoel en functie

Het gebruiksdoel van de SVT Spelling is driedig:

- De toets kan worden ingezet voor niveaubepaling van de leerlingen, zowel individueel als per groep.
- De toets kan worden ingezet voor de voortgangsbepaling van de leerlingen.
- De toets kan worden ingezet voor een foutenanalyse voor de zwakke spellers.

Inhoudelijke theoretische inkadering:

Bij het samenstellen van de dictees is met inachtneming van de geldende referentieniveaus geprobeerd de spellingsproblemen die zich voor doen in de Nederlandse taal zo goed mogelijk te representeren, met gangbare woorden die worden beschouwd als behorend tot de actieve woordenschat die kinderen zouden moeten beheersen.

Naast het Referentiekader Taal wordt ook gerefereerd aan de tussendoelen voor groep 4 en 5 en groep 6 tot en met 8 die zijn opgesteld door het Expertisecentrum Nederlands.

Inhoud van het toetspakket

Het toetspakket bestaat uit:

- Handleiding en Verantwoording
- Handleiding en Verantwoording Aanvulling
- Instructieboekje

- 1 antwoordblad voor SVT Spelling Woorden
- 3 antwoordbladen voor SVT Spelling Werkwoorden
 - Blok 1, versie A + B
 - Blok 2, versie A + B
 - Blok 3, versie A + B
- 10 normtabellen SVT Spelling Woorden
 - Blok 1 en 2 A
 - Blok 1 en 2 B
 - Blok 2 en 3 A
 - Blok 2 en 3 B
 - Blok 3 en 4 A
 - Blok 3 en 4 B
 - Blok 4 en 5 A
 - Blok 4 en 5 B
 - Blok 5 en 6 A
 - Blok 5 en 6 B
- 4 normtabellen SVT Spelling Werkwoorden
 - Blok 1 en 2 A
 - Blok 1 en 2 B
 - Blok 2 en 3 A
 - Blok 2 en 3 B
- Voorbeeld Individueel Rapport
- Informatiebrochure voor ouders
- Toegangscode voor de scorings- en normeringsapplicatie

2. Beoordeling van de kwaliteitsaspecten

De beoordeling vindt plaats volgens het 'Beoordelingskader voor de psychometrische aspecten van (reeksen van) toetsen uit leerlingvolgsystemen (LOVS)', 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), Dr. Desiree Joosten-Ten Brinke (onderwijskundig expert) en mevrouw Paulyn K. Berding-Oldersma MSc (secretaris).

De kwaliteit van de steekproef

S1.1. Is de steekproef representatief?

Bevindingen:

De normering van de toets vond plaats in de maanden maart en november 2013. De normeringssteekproef in het voorjaar bestond uit 18 basisscholen en in het najaar uit 17 basisscholen. Bij 8 basisscholen uit de eerste normeringssteekproef is bij de leerlingen zowel versie A als versie B afgenomen, en eveneens bij 8 basisscholen uit de tweede normeringsteekproef. Bij in totaal 2316 leerlingen is zowel versie A als versie B afgenomen van de SVT Spelling Woorden en bij 1238 leerlingen is zowel versie A als versie B afgenomen van de SVT Spelling Werkwoorden. De toetsresultaten en andere gegevens van voornoemde leerlingen vormen de kalibratiesteekproef (zie S1.2).

Tabel 5.1 bevat de gegevens van de normeringssteekproef SVT Spelling Woorden en Werkwoorden. Bij versie A van SVT Spelling Woorden werden gegevens van 1864 leerlingen afkomstig van 11 scholen verzameld en bij versie B gegevens van 1389

leerlingen afkomstig van 8 scholen. Bij versie A en B van SVT Spelling Werkwoorden werden gegevens van 2316 leerlingen afkomstig van 35 scholen verzameld. In totaal namen 5569 leerlingen deel aan de normeringssteekproef.

De representativiteit van de steekproef is onderzocht met betrekking tot regio, stedelijkheid, sekse, herkomst en leerlinggewicht.

Wat betreft regio werden de leerlingen onderverdeeld in de vier regio's die door het CBS gehanteerd worden. Bij versie A is er enige mate een oververtegenwoordiging (5 %) van de regio Oost en een ondervertegenwoordiging (5%) van de regio Zuid, terwijl er bij versie B een oververtegenwoordiging was van de regio Noord (5%) en de regio Oost (9%) en een ondervertegenwoordiging van de regio West (9%) en de regio Zuid (4%).

Wat betreft stedelijkheid werden de leerlingen onderverdeeld in de vijf gradaties van stedelijkheid die door het CBS onderscheiden worden. Zowel bij voornoemde verdeling als bij de verdeling stedelijk versus platteland zien we afwijkingen van 6% tot 10% ten opzichte van de verdeling van de landelijke populatie.

Wat betreft sekse is de verdeling in de steekproef nagenoeg gelijk aan die aan de verdeling in de landelijke populatie.

Wat betreft herkomst is uitgegaan van de door het CBS gehanteerde definitie: allochtone leerlingen zijn leerlingen van wie ten minste één van de ouders niet in Nederland is geboren. Het totale percentage van leerlingen met een allochtone achtergrond is 27,7% (versie A) en 27,2% (versie B), terwijl het landelijk percentage 23,7% bedraagt. In de steekproeven ligt het percentage leerlingen met een allochtone achtergrond dus ongeveer 4% hoger.

Wat betreft de leerlinggewichten zijn 0,0 leerlingen ongeveer 5% ondervertegenwoordigd en de 0,3 en 1,2 leerlingen ongeveer 5% oververtegenwoordigd.

Om de normeringssteekproef in overeenstemming te brengen met de verdeling in de populatie is binnen de normeringssteekproeven gewogen voor regio en herkomst. Bijlage VII laat zien dat na weging de verdeling van de achtergrondkenmerken zeer dicht bij de verdeling in de populatie liggen. Een uitzondering vormt de mate van stedelijkheid voor de steekproef van versie B want die is ook na weging nog enigszins oververtegenwoordigd.

Conclusie:

De representativiteit van de steekproef is onderzocht met betrekking tot regio, urbanisatiegraad, en sekse, herkomst en leerlinggewicht. Om verschillen tussen de steekproeven en de landelijk verdeling te minimaliseren is met succes weging toegepast. Op aspect S1.1 wordt aan SVT Spelling het oordeel '**goed**' toegekend.

S1.2. In geval van een onvolledig dataverzamelingsdesign: is het design adequaat?

Bevindingen:

In tabel 5.4 van de Handleiding en Verantwoording staat het normeringsschema voor de kalibratie van de SVT Spelling Woorden en de SVT Spelling Werkwoorden. Volgens het normeringsschema zijn er 11 normeringsmomenten, te beginnen met groep na 7 maanden onderwijs (DL = 7) en eindigend met groep 8 na 57 maanden onderwijs (DL = 57). Het normeringsschema voor SVT Spelling Woorden ziet er als volgt uit:

- op normeringsmoment 1 (DL = 7) maakt groep 3 blok 1 & 2 (= 30 items);
- op normeringsmoment 2 (DL = 13) maakt groep 4 blok 1 & 2 en blok 2 & 3;
- op normeringsmoment 3 (DL = 17) maakt groep 4 blok 1 & 2 en blok 2 & 3;
- op normeringsmoment 4 (DL = 23) maakt groep 5 blok 1 & 2, blok 2 & 3 en blok 3 & 4;
- op normeringsmoment 5 (DL = 27) maakt groep 5 blok 1 & 2, blok 2 & 3 en blok 3 & 4;
- op normeringsmoment 6 (DL = 33) maakt groep 6 blok 2 & 3, blok 3 & 4 en blok 4 & 5;
- op normeringsmoment 7 (DL = 37) maakt groep 6 blok 2 & 3, blok 3 & 4 en blok 4 & 5;
- op normeringsmoment 8 (DL = 43) maakt groep 7 blok 3 & 4, blok 4 & 5 en blok 5 & 6;
- op normeringsmoment 9 (DL = 47) maakt groep 7 blok 3 & 4, blok 4 & 5 en blok 5 & 6;
- op normeringsmoment 10 (DL = 53) maakt groep 8 blok 3 & 4, blok 4 & 5 en blok 5 & 6;
- op normeringsmoment 11 (DL = 57) maakt groep 8 blok 3 & 4, blok 4 & 5 en blok 5 & 6;

Het normeringsschema voor Werkwoorden is dat op de normeringsmomenten 6 t/m 11 alle leerlingen uit de bovenbouw (groep 6, 7 en 8) blok 1 & 2 en blok 2 & 3 maken.

Het normeringsschema of kalibratiedesign laat zien dat er sprake is van voldoende linking. Het meetmodel dat voor de kalibratie gebruikt is, is het Rasch-model. Wat betreft de nauwkeurigheid van de parameterschattingen kan geconcludeerd worden dat deze voldoen aan het COTAN criterium dat stelt dat de standaardfout van de moeilijkheidsparameter niet groter moet zijn dan de standaarddeviatie van de vaardigheidsparameter maal een constante c waarbij $c < 0,2$ als 'goed' aangemerkt wordt. Bijlage III laat zien dat de items aan dit criterium voldoen. In tabel 4.6 van de Handleiding en Verantwoording worden de statistieken van de Rasch-schalen voor de blokken 1 & 2, 2 & 3, 3 & 4, 5 & 6 voor de versies A en B van personen en items gepresenteerd: SD theta, RMSE (Root Mean Square Error), separatie-index (items) en betrouwbaarheid (personen). Op basis van de waarden van voornoemde statistieken kan geconcludeerd worden dat het Rasch-model past. Ter onderbouwing van die conclusie is ook gekeken naar de passing van de items. De infit- en outfit statistieken van de items bevestigen de conclusie van voldoende passing.

Conclusie:

Het dataverzamelingsdesign is adequaat. Op aspect S1.2 wordt aan de toetsen SVT Spelling Woorden en SVT Spelling Werkwoorden het oordeel '**goed**' toegekend.

Normering

N1.2.1. Zijn de normgroepen groot genoeg?

Bevindingen:

In tabel 5.5 en tabel 5.6 worden de aantallen leerlingen vermeld voor de normering van SVT Spelling Woorden en SVT Spelling Werkwoorden voor versie A en B. Er zijn normtabellen voor 11 normeringsmomenten (zie S1.2) voor vijf toetsen (blok 1 & 2, blok 2 & 3 etc.). Voor elk normeringsmoment en elke toets zijn de toetsgegevens van 299 tot 431 leerlingen gebruikt.

Conclusie:

Uitgaande van de COTAN beoordelingscriteria wordt op aspect N1.2.1 aan de toetsen SVT Spelling Woorden en SVT Spelling Werkwoorden het oordeel '**goed**' toegekend.

N1.2.2. Zijn de normgroepen representatief?

Bevindingen:

De representativiteit van de steekproeven werd in S1.1 2.1 besproken en daar werd geconstateerd dat de steekproeven representatief waren voor regio, stedelijkheid, sekse, herkomst en leerlinggewicht.

Conclusie:

Op aspect N1.2.1 wordt aan de toetsen SVT Spelling Woorden en SVT Spelling Werkwoorden het oordeel '**goed**' toegekend.

Betrouwbaarheid

B1.1. Zijn of worden de betrouwbaarheidsgegevens correct berekend?

Bevindingen:

Gebruikmakend van het kader van het Rasch-model geeft de betrouwbaarheid voor personen de rangschikking aan van personen op de latente trek. De betrouwbaarheid wordt berekend als de verhouding van de ware scorevariantie op de latente trek en de geobserveerde variantie op de latente trek. Deze coëfficiënt wordt in de literatuur onder andere beschreven in het boek Measurement Essentials van Wright en Stone (1999) en vertoont qua interpretatie grote overeenkomst met de betrouwbaarheidscoëfficiënt uit de klassieke testtheorie. Daarnaast is Cronbach's alfa, de parallelle betrouwbaarheid en de test-her-test betrouwbaarheid berekend. De laatste berekening is echter op zeer weinig leerlingen gebaseerd.

Naast voornoemde betrouwbaarheidscoëfficiënten zijn ook accuratesse-indices berekend. Daaronder wordt verstaan de mate waarin de leerling wordt geclassificeerd in de categorie waarin hij/zij behoort op basis van de werkelijke vaardigheidsscore. Omdat de SVT Spelling ook bedoeld is om het voortgangsniveau van de leerling te bepalen, is het belangrijk om te weten hoe betrouwbaar de classificaties van de niveau-indelingen zijn. De indelingen zijn de zogenaamde Cito indelingen van A t/m E en I t/m V. In de Bijlage van de Aanvulling bij de Handleiding en Verantwoording worden per didactische leermaand (DL) en voor de versies A en B de marginale accuratesse-indices gepresenteerd voor alle grenswaarden (grens E-D, grens D-C, grens C-B, grens B-A; V-IV, IV - III, III - II en II - 1) en voor de classificering waarbij alle grenswaarden tegelijk worden toegepast. In dit verband dient opgemerkt te worden dat het bij de beoordeling van de kwaliteit van de accuratesse-indices zou helpen als er een goede toelichting met voorbeeld(en) voorhanden was van hoe de indices berekend en geïnterpreteerd moeten worden.

Conclusie:

De betrouwbaarheidsgegevens worden correct berekend en op aspect B1.1 wordt aan de toetsen SVT Spelling het oordeel '**goed**' toegekend.

B1.2. Zijn de betrouwbaarheidsgegevens voldoende gezien de beslissingen die met de toets genomen worden?

Bevindingen:

De betrouwbaarheden van de vijf toetsen SVT Spelling Woorden versies A en B (blok 1 & 2, blok 2 & 3, blok 3 & 4, blok 4 & 5 en blok 5 & 6) die berekend zijn op basis van het Rasch-model zijn 0,82 of hoger. De betrouwbaarheden van de twee toetsen SVT Spelling Werkwoorden versies A en B (blok 1 & 2 en blok 2 & 3) waren 0,72 of hoger.

Cronbach's alfa is berekend voor versie A en B van voornoemde toetsen voor Spelling Woorden en Spelling Werkwoorden, voor onderscheiden groepen en voor de totale groep. De betrouwbaarheden volgens Cronbach's alfa waren 0,74 of hoger.

Onderzoek naar de paralleliteit van versie A en B bevestigde dat de versies parallel waren. Daardoor kon voor kalibratiesteekproef I waarbij leerlingen beide versies maakten, de parallelbetrouwbaarheid berekend worden. De betrouwbaarheden van voornoemde toetsen voor Spelling Woorden en Spelling Werkwoorden, waren voor onderscheiden groepen en voor de totale groep 0,78 of hoger.

Ook de test-hertest betrouwbaarheden waren hoger dan 0,72 maar waren berekend op zeer weinig leerlingen van een zogenoemde 'zwarte' school in Amsterdam.

De accuratesse-indices voor de grenswaarden zijn zeer hoog, gemiddeld ongeveer 0,90, terwijl die voor 'alle grenzen tegelijk' aanzienlijk lager zijn, gemiddeld ongeveer 0,65.

Conclusie:

Met verwijzing naar het beoordelingssysteem van de COTAN kan, gegeven de gerealiseerde betrouwbaarheidscoëfficiënten en accuratesse-indices, op aspect B1.2 het oordeel 'goed' toegekend worden aan de toetsen SVT Spelling Woorden en SVT Spelling Werkwoorden.

Validiteit

V1. Dragen de items in de toets bij aan de validiteit van de toets (hierbij gaat het om aspecten als relevantie, objectiviteit en efficiëntie van de items)

Bevindingen:

De toetsontwikkelaars hebben de keuze gemaakt om opgaven rond spelling van woorden en spelling van werkwoorden van elkaar te scheiden. Die keuzes scheppen duidelijkheid en vergemakkelijken het scoren en interpreteren van de toetsen.

De vorm van telkens twee blokken met opgaven in twee versies geeft voldoende mogelijkheid tot planning van afname eenmaal of tweemaal per jaar. De belasting van leerkrachten en leerlingen door het jaar is hiermee in verhouding tot de opbrengst.

In de inhoudelijke verantwoording wordt duidelijk dat de ontwikkelaars zich richten op de in het Referentiekader Taal vastgelegde uitwerkingen van het domein Spelling, met gebruikmaking van een aantal passende aanvullende bronnen om de doorgaande lijn in ontwikkeling van leerlingen op dit vlak te volgen. Dit levert een relevante verdeling en opbouw van subvaardigheden over toetsonderdelen op.

Tevens is gepast aandacht gegeven aan onderbouwing van de keuze voor een dictee in alle gevallen, en geen gebruikmaking van meerkeuzeopgaven. Hierbij zou wellicht nog een onderbouwing hebben gepast van de keuze om het dictee door de leerkracht zelf te laten voorlezen – een ingesproken cd met de dictees zal de betrouwbaarheid van de afname verhogen.

De opgaven en de vraagstelling zijn doorgaans helder en eenduidig, slechts over een enkele opgave zou discussie tussen deskundigen kunnen ontstaan.

Merk op dat de beoordeling van dit aspect zich hieronder beperkt tot het statistisch/psychometrisch onderzoek dat verricht is.

Hoofdstuk 7 van de Handleiding en Verantwoording betreft verslag van onderzoek naar de begripsvaliditeit van de toetsen SVT Spelling Woorden en SVT Spelling Werkwoorden. Het onderzoek heeft betrekking op:

- Einddimensionaliteit en psychometrische kwaliteit van de items. Voor dit onderzoek wordt verwezen naar S1.2 waar de passing van het Rasch-model besproken werd en naar V1.1 waar de factoriële structuur van de SVT Spelling onderzocht werd. Daarnaast zijn er ook netwerkanalyses uitgevoerd om de onderlinge verbanden tussen items te onderzoeken. Uit de netwerkanalyses bleek dat er geen duidelijk te onderscheiden clusters van items waren en er dus geen aanleiding was meerdimensionaliteit te veronderstellen en dat de weinige items die buiten het netwerk vaak ook een hoge outfit in de passing van het Rasch-model hadden en zogenoemde 'instinkers' waren.

- Ontwikkeling van de spellingsvaardigheid. Wanneer men naar de vaardigheidsscores op de onderscheiden normeringsmomenten kijkt, blijkt dat spellingsvaardigheid zich op een lineaire wijze ontwikkelt. Vanaf midden groep 3 leren leerlingen spellen en de spellingsvaardigheid blijft tot en met groep 8 toenemen.

- Vraagonzuiverheid. Onderzocht is vraagonzuiverheid ten opzichte van sekse en etnische achtergrond. Het effect van de weinige vraagonzuivere items op de gemiddelde itemscores van de focus- en referentiegroepen bleek gering.

- Groepsverschillen. Op de spellingtoetsen blijken meisjes over het algemeen hoger te scoren dan jongens. Verrassend was dat er geen significante verschillen tussen leerlingen met leerlinggewicht 0 en 0,3 gevonden werden maar zoals verwacht waren de verschillen met leerlingen met leerlinggewicht 1,2 wel significant.

- Onderzoek naar convergente en divergente validiteit. Convergente validiteit van SVT Spelling Woorden is onderzocht met Cito Spelling en het PI dictee door de correlaties tussen de toetsen te berekenen en de normscores te vergelijken. De correlaties met Cito Spelling zijn, met uitzondering van groep 3, hoog te noemen. Daarnaast zijn de normscores, dwz de niveau-indeling A t/m E, van Cito Spelling en SVT Spelling Woorden vergeleken. In de kruistabellen van de twee versies, ziet men de mate van overeenstemming tussen beide niveau-indelingen. Hoewel er de nodige discrepanties tussen de indelingen bestaan, blijkt de hoogte van coëfficiënt gamma op een matig tot sterk verband te wijzen. De correlaties tussen SVT Spelling Woorden en het PI-dictee zijn hoog. Om de divergente validiteit te onderzoeken is naar de relatie tussen de SVT Spelling Woorden en SVT Spelling Werkwoorden en Cito Begrijpend Lezen gekeken. Zoals verwacht blijkt uit de correlaties een zeer zwak verband.

Conclusie:

Op aspect V1.1 wordt aan de toetsen SVT Spelling Woorden en SVT Spelling Werkwoorden het oordeel '**goed**' toegekend.

Het volg-aspect

VA1.1. Is er een voldoende empirische onderbouwing van de schaal waarop de groei van een leerling wordt uitgedrukt? Wordt groei op een adequate manier gemeten?

Bevindingen:

In paragraaf 4.3 van de Handleiding en Verantwoording wordt de kalibratiestudie beschreven. De kalibratiestudie bestond uit twee stappen. In de eerste stap werd met

behulp van een non-lineaire factoranalyse (NOHARM) de aanname van eendimensionaliteit onderzocht. NOHARM berekent de root mean square (RMSR) van de covarianties als een globale passingsmaat van het model voor de data bij het gegeven aantal dimensies. De analyse van de resultaten in tabel 4.1 voor versie A en B van SVT Spelling Woorden en van SVT Spelling Werkwoorden laat zien dat de passingsmaten voor versie A en B vergelijkbaar zijn en dat hoewel de passing van het tweedimensionale model beter is dan die van het eendimensionale model, de verschillen dusdanig klein zijn dat de aanname van eendimensionaliteit gerechtvaardigd is. In de tweede stap werd de passing van het Rasch-model onderzocht. In S1.2 werden statistieken over de passing gepresenteerd en werd geconcludeerd dat de aanname van passing van het eendimensionale Rasch-model gerechtvaardigd is.

Het voorgaande betekent dat men beschikt over één schaal voor SVT Spelling Woorden en één schaal voor SVT Spelling Werkwoorden waarop de items van de onderscheiden toetsen (en leerlingen) op afgebeeld kunnen. Bedoelde schalen maken het mogelijk om zowel het niveau als de groei van de leerlingen vast te stellen.

Voor de vraag of groei op een adequate manier gemeten wordt, wordt verwezen naar het 'individueel rapport' en de normtabellen. In het 'individueel rapport' worden onder de normaalverdeling de normscores grafisch weergegeven. In het rapport op pagina 40 van de Handleiding en Verantwoording wordt de toets vermeld die een leerling gemaakt heeft, in dit geval blok 3 & 4 Woorden, op normeringsmoment DL = 28, de behaalde ruwe score, de percentielscore, de twee niveau-indelingen, de DLE, de vaardigheidsscore met bijbehorend 90% betrouwbaarheidsinterval. Door de vaardigheidsscore van de leerling af te zetten tegen eerdere toetsresultaten kan men constateren of er sprake is van vooruitgang. Als de betrouwbaarheidsintervallen van de laatste en de voorgaande toetsafname elkaar niet overlappen, kan men bijna zeker concluderen dat er sprake is van werkelijke vooruitgang. Daarnaast kan men ook percentielscores, de gebruikelijke twee niveau-indelingen en DLE scores van opeenvolgende toetsafnames vergelijken om groei vast te stellen (zie ook VA1.2)

Conclusie:

Op aspect VA1.1 wordt aan de toetsen SVT Spelling Woorden en SVT Spelling Werkwoorden het oordeel '**goed**' toegekend.

VA1.2. Worden er gegevens verstrekt over hoe groei geïnterpreteerd dient te worden? Wordt de betrouwbaarheid van de groei op die schaal adequaat weergegeven?

Bevindingen:

In hoofdstuk 1 van de Aanvulling bij de Handleiding en Verantwoording worden drie aspecten mbt het volgen van leerlingen onderscheiden:

1. Inhoudelijke analyse. De voornaamste instrumenten daarvoor zijn het foutenoverzicht en de foutenanalyse;
2. De vaardigheidsscore. De toetsscores van de leerling zijn op een vaardigheidsschaal geplaatst aan de hand waarvan de ontwikkeling gevolgd kan worden. Voor een interpretatie van de vaardigheidsscore kan de gemiddelde groei van de vaardigheidsscore behulpzaam zijn. Met behulp van de betrouwbaarheidsintervallen van opeenvolgende vaardigheidsscores kan men vaststellen of er sprake is van werkelijke groei. Het blijkt dat bij SVT Spelling Woorden pas na een half jaar werkelijke groei geconstateerd kan worden en bij SVT Spelling Werkwoorden pas na een jaar.

3. De normscore. Door normscores van opeenvolgende toetsafnames te vergelijken, kan de ontwikkeling van een leerling in beeld gebracht worden. Hierbij kan men gebruik maken van een percentielscore, een percentielscore in niveaus zoals de twee gebruikelijke niveau-indelingen of een DLE-score.

Aanbevolen wordt om de leervorderingen van leerlingen bij te houden aan de hand van een vaardigheidsscore in combinatie met een normscore.

Conclusie:

Op aspect VA1.2 wordt aan de toetsen SVT Spelling Woorden en SVT Spelling Werkwoorden het oordeel '**goed**' toegekend.

Inzicht in leervorderingen

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

Bevindingen:

De toetsen SVT Spelling Woorden en SVT Spelling Werkwoorden gaan vergezeld van een informatiebrochure voor ouders getiteld 'De SVT Spelling en uw kind'. In die brochure wordt aan ouders uitgelegd hoe de spellingsvaardigheid van hun kind getoetst wordt en hoe de prestatie van hun kind in cijfers of letters uitgedrukt wordt. Achtereenvolgens komen in de brochure de volgende onderwerpen aan de orde:

- Het leerlingvolgsysteem: verschil tussen niet-methodegebonden en methodegebonden toetsen; makkelijke en moeilijke opgaven; een- of tweemaal per jaar afnemen;
- Inhoud van de SVT Spelling: spellingsvaardigheid wordt met een dictee getoetst; SVT Spelling bestaat uit een reeks van deeltaetsen, zogenoemde 'blokken';
- De interpretatie van de score: percentielscore en niveau-indelingen; de DLE-systematiek; de vaardigheidsscore;
- Het individuele rapport: opbouw van het rapport; het rapport van Albert Alberts; percentielscore en normaalverdeling; DLE; vaardigheidsscore en 90%-betrouwbaarheidsinterval;
- De SVT Spelling en het leerlingadministratiesysteem: leergroei grafisch in beeld brengen.

De terugkoppeling is nu alleen (nog) direct helder en inzichtelijk voor de leerkracht en andere professionals in de school. Voor gebruik door ouders/verzorgers is meer toelichting en ook in andere vormen qua taalgebruik en visuele ondersteuning nodig. De gebruikte termen en beelden zijn voor hen te weinig herkenbaar en lastig tot niet te interpreteren. De leerkracht zal een belangrijke rol moeten spelen voordat ouders zelfstandig deze uitkomsten kunnen interpreteren.

Conclusie:

Op aspect I1.1 wordt aan de toetsen SVT Spelling Woorden en SVT Spelling Werkwoorden het oordeel '**voldoende**' toegekend.

3. Verzamelstaat

Kwaliteitsaspect	Code	Oordeel
De kwaliteit van de steekproef	S1.1	Goed
	S1.2	Goed
Normering	N1.1	Goed
	N1.2	Goed
Betrouwbaarheid	B1.1	Goed
	B1.2	Goed
Validiteit	V1.1	Goed
	VA1.1	Goed
Volg-aspect	VA1.2	Goed
	I1.1	Voldoende

4. Literatuurlijst

- Braams, T. & Vos T. de (2015). *Schoolvaardigheidstoets Spelling. Handleiding en Verantwoording*. Amsterdam: Boom test uitgevers.
- Braams, T. & Vos T. de (2015). *Schoolvaardigheidstoets Spelling. Handleiding en Verantwoording, aanvulling*. Amsterdam: Boom test uitgevers.
- Braams, T. & Vos T. de (2015). *Schoolvaardigheidstoets Spelling*. Amsterdam: Boom test uitgevers.