

1. Uitgangspunten van de toetsconstructie

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 wetenschappelijke verantwoording heeft betrekking op de uitgangspunten van de toetsconstructie, de normen, de betrouwbaarheid en meetnauwkeurigheid en de validiteit. De Handleiding heeft betrekking op het gebruik van de toets, communicatie over de toetsgegevens en de inhoudsverantwoording.

Algemeen

Het Cito Volgsysteem primair en speciaal onderwijs beoogt de vorderingen van individuele leerlingen, groepen leerlingen en het onderwijs op school van groep 1 tot en met groep 8 te volgen en te evalueren. De toetsen Spelling 3.0 groep 6 zijn een onderdeel van het Cito Volgsysteem primair en speciaal onderwijs en zijn bedoeld voor leerlingen in groep 6 van het primair en speciaal onderwijs.

De toetsen voor groep 3 tot en met groep 8 worden vanaf het schooljaar 2014/2015 geleidelijk vervangen door de nieuwe toetsen Spelling 3.0. De toetsen voor groep 6 vormen dan samen de LVS toetsen Spelling 3.0 groepen 3, 4, 5 en de groepen 7 en 8 een systeem om vast te stellen hoe goed leerlingen kunnen spellen en hoe hun vaardigheid zich in de basisschoolperiode ontwikkelt. Onderstaande beschrijving is gebaseerd op de Handleiding.

Meetpretentie

De toetsen in de toets pakketten Spelling 3.0 van het Cito Volgsysteem primair en speciaal onderwijs zijn bedoeld om vast te stellen hoe goed een leerling kan spellen en hoe de spellingvaardigheid van de leerling zich in de loop van de jaren ontwikkelt.

Het vaststellen van de spellingvaardigheid gebeurt door de leerling woorden te laten opschrijven (actieve spelling). De spellingregels zelf worden niet expliciet bevraagd. De leerling laat indirect zien dat hij of zij de spellingregels beheerst door de gevraagde woorden correct te schrijven.

Doelgroep

De toetsen Spelling 3.0 groep 6 zijn bedoeld voor leerlingen in groep 6 van het primair en speciaal onderwijs, maar kunnen ook gebruikt worden voor leerlingen uit andere jaargroepen die werken op het niveau van groep 6 en voor leerlingen met een ontwikkelingsachterstand en/of extra onderwijsbehoeften. Voor deze groepen speciale leerlingen zijn geen afzonderlijke normen vastgesteld en de toetsresultaten van deze leerlingen worden geïnterpreteerd met behulp van de gemiddelde vaardigheidsscores voor leerlingen uit het regulier onderwijs.

Gebruiksdoel en functie

Doel van de toetsen Spelling 3.0 is het in kaart brengen van het vaardigheidsniveau en de ontwikkeling van de leerlingen op het gebied van de spelling, voor groep 6 specifiek het schrijven van niet-klankzuivere woorden met twee of meer lettergrepen. De toetsen maken het mogelijk om:

- De vaardigheid spelling van zowel individuele leerlingen als groepen leerlingen (groeps- en schoolniveau) te beoordelen via een vergelijking van de behaalde scores met de scores van een landelijke referentiegroep oftewel niveaubepaling.

- De ontwikkeling van de vaardigheid spelling van zowel individuele leerlingen als groepen leerlingen (groeps- en schoolniveau) door de leerjaren heen te volgen oftewel progressiebepaling.

Inhoudelijke theoretische inkadering:

De inhoud van de toetsen Spelling 3.0 is gebaseerd op het domein Begrippenlijst en Taalverzorging, en dan het onderdeel Taalverzorging, beschreven in het Referentiekader Taal en Rekenen. De toetsen sluiten aan bij de indeling die is gehanteerd in het Referentiekader Taal. In de publicatie 'Leerstoflijnen begrippenlijst en taalverzorging beschreven' van de SLO, is aangegeven hoe de opbouw van de leerstoflijnen eruit kan zien voor de verschillende groepen. Voor de inhoud van de toetsen zijn deze uitwerkingen bepalend geweest voor de theoretische basis als voor de indeling van het categorieënoverzicht. Verfijning van de categorieën is mede gebaseerd op basis van analyse van methoden (taal) die veel gebruikt worden in het basisonderwijs.

De koppeling die gemaakt wordt met het Referentiekader levert voor scholen een meerwaarde.

Inhoud van het toetspakket

Het toetspakket Spelling 3.0 groep 6 bestaat uit de volgende documenten:

- Handleiding, deze bevat informatie over:
 - de afname van de toets (hfdst. 2),
 - nakijken en verwerken van toetsgegevens (hfdst. 3),
 - interpretatie van de toetsresultaten op leerling- en groepsniveaus (hfdst 4),
 - algemene aandachtspunten voor het schoolplan (hfdst 5),
 - inhoudsverantwoording (hfdst 6),
 - communiceren over toetsresultaten met leerling en ouders (hfdst 7),
 - achtergrondinformatie en veel gestelde vragen (hfdst 8) en
 - enkele bijlagen
- Vier toetsen:
 - Toets E5M6 (makkelijke variant van de toets M6)
 - Toets M6 (Medio groep 6)
 - Toets E6 (Eind groep 6)
 - Toets M6E6 (makkelijke variant van de toets E6)
- Afnamekaarten met aanwijzingen voor de papieren afname van de toetsen
- Nakijkaarten
- Antwoordbladen
- Tabellen voor de vier toetsen voor het bepalen van de vaardigheidsscore en – niveau
- Wetenschappelijke verantwoording.

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:

Sinds 2013/2014 wordt door Cito een nieuwe werkwijze voor het normeren van leerlingvolg-systemen toegepast. Een belangrijk aspect van die werkwijze is om de normeringssteekproef (ongeveer 50 procent) te baseren op gegevens verkregen via een embedded field onderzoek en (ongeveer 50 procent) op gegevens die afkomstig zijn van de tweede generatie LVS-toetsen via Cito dataretour. Bij de normering van de derde generatie LVS-toetsen is rekening gehouden met de variabelen regio, urbanisatiegraad, schooltype (stratum) en sekse, waarbij de eerste drie variabelen op schoolniveau zijn gedefinieerd en de laatste variabele op leerlingniveau. Op pagina 46 en 47 van de wetenschappelijke verantwoording wordt een algoritme beschreven dat een representatieve steekproef moet garanderen. Niettemin werd er naast die garantie een controle op de representativiteit uitgevoerd door de populatieverdelingen van gegevens uit DUO te vergelijken met de steekproefverdelingen van M6 en E6. In de tabellen 4.5 tot en met 4.8 worden de resultaten van de representativiteitsanalyses gepresenteerd. De effectgroottes zijn zeer klein en de conclusie is dan ook dat de normeringssteekproeven een zeer goede afspiegeling van de populatieverdelingen zijn.

In tabel 4.4 van de wetenschappelijke verantwoording worden de aantallen leerlingen gegeven die meegenomen zijn bij de normering. Voor M6 betreft dat 3053 leerlingen (1432 leerlingen uit de steekproef en 1621 leerlingen via dataretour) afkomstig van 138 scholen. Voor E6 betreft dat 2836 leerlingen (1180 leerlingen uit de steekproef en 1656 leerlingen via dataretour) afkomstig van 132 scholen.

Conclusie:

De representativiteit van de steekproef is onderzocht met betrekking tot regio, urbanisatiegraad, schooltype en sekse. Op aspect S1.1 wordt aan de toetsen Spelling 3.0 groep 6 het oordeel '**voldoende**' toegekend.

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

Bevindingen:

Bij de ontwikkeling van de toetsen Spelling 3.0 groep 6 is geprobeerd om door middel van embedded field designs bias in de normen te vermijden door de afnamesituatie waarin de toets wordt afgenomen tijdens het normeringsonderzoek zoveel mogelijk te laten lijken op de afnamesituatie na uitgave van de toets.

Het normeringsonderzoek gaat uit van een onvolledig dataverzamelingsdesign. In totaal lieten scholen hun leerlingen drie taken maken: één taak uit de tweede generatie van de LVS-toetsen en twee taken met nieuw materiaal voor de derde generatie. Deze drie taken zaten samen in een enkel boekje en telden alle drie mee voor het resultaat van de leerling. In figuur 4.3 en 4.4 worden de embedded field designs weergegeven voor respectievelijk M6 en E6. In de designs zien we dat er voor M6 en E6 zeven toetsversies afgenomen zijn. Toetsversie 1 van M6 bestaat uit een taak uit M6 LVS 2e generatie, een taak E5M6 deel 1 en een taak E5M6 deel 2. Toetsversie 2 bestaat uit een taak M6 LVS 2e

generatie, een taak M6 deel 1 en een taak M6 deel 2, etc.. Behalve de toetsen M6 en E6 worden ook tussentoetsen, reservetaken en bij design E6 ook een taak M7 afgenomen. Een enkele toetsversie werd bij tussen de 168 en 305 leerlingen afgenomen.

In het normeringsonderzoek M6 in 2015 zijn 205 verschillende items voorgelegd aan 1530 leerlingen van groep 6. Hiervan hebben 1488 leerlingen de volledige taken gemaakt en die leerlingen zijn verdeeld over zeven boekjes. Elk boekje bestond uit 80 opgaven, verdeeld over drie taken. De 30 opgaven uit de bestaande taak Spelling werd door alle 1488 leerlingen uit het normeringsonderzoek gemaakt. De overige opgaven kwamen steeds in twee boekjes voor en werden gemiddeld door 435 leerlingen gemaakt.

In het normeringsonderzoek E6 in 2015 zijn 205 verschillende items voorgelegd aan 1430 leerlingen van groep 6. Hiervan hebben 1413 leerlingen de volledige taken gemaakt en die leerlingen zijn verdeeld over zeven boekjes. Elk boekje bestond uit 75, 80 of 85 opgaven, verdeeld over drie taken. De 30 opgaven uit de bestaande taak Spelling werd door alle 1413 leerlingen uit het normeringsonderzoek gemaakt. De overige opgaven kwamen steeds in twee boekjes voor en werden gemiddeld door 409 leerlingen gemaakt.

Uit het kalibratieonderzoek blijkt dat voor de vier (tussen)toetsen zowel de items, gegeven de M- en S-toetsen, als de gehele toetsen, gegeven de R1c-toetsen, passen bij het itemresponstheoriemodel OPLM. Dit betekent dat er sprake is van een eendimensionale vaardigheidsschaal waar items en leerlingen op afgebeeld kunnen worden. Op grond van de gegevens uit de kalibratie zijn de definitieve selecties van items gemaakt voor de uitgaven van de toetsen Spelling 3.0 M6 en Spelling 3.0 E6.

Conclusie:

Het onvolledige dataverzamelingsdesign is adequaat. Op aspect S1.2 wordt aan de toetsen Spelling 3.0 groep 6 het oordeel '**voldoende**' toegekend.

Normering

N1.2.1. Zijn de normgroepen groot genoeg?

Bevindingen:

In tabel 4.9 van de wetenschappelijke verantwoording wordt voor M6 en E6 de normtabel met relatieve normen op leerlingniveau gepresenteerd. Voor aantallen leerlingen en scholen zie S1.1. Voor beide afnamemomenten worden vaardigheidsverdelingen gepresenteerd, d.w.z. gemiddelde score, standaarddeviatie en de percentielen P10, P20, P25, P40, P50, P60, P75, P80, P90. Met behulp van deze percentielen kunnen de twee niveau-indelingen (A t/m E en I t/m V) opgesteld worden.

In tabel 4.11 van de wetenschappelijke verantwoording wordt voor M6 en E6 de normtabel met relatieve normen op schoolniveau gepresenteerd. Voor beide afnamemomenten worden vaardigheidsverdelingen gepresenteerd, d.w.z. gemiddelde score, standaarddeviatie en de percentielen P10, P20, P25, P40, P50, P60, P75, P80, P90.

Conclusie:

De normgroepen zijn groot genoeg (zie S1.1). Op aspect N1.2.1 wordt aan de toetsen Spelling 3.0 groep 6 het oordeel '**voldoende**' toegekend.

N1.2.2. Zijn de normgroepen representatief?

Bevindingen:

De representativiteit van de steekproeven werd in S1.1 en 2.1 besproken en daar werd geconstateerd dat de steekproeven representatief waren voor regio, urbanisatiegraad, schooltype en sekse. Voor Spelling 3.0 groep 6 geldt dat de normen geldig zijn tot en met 2025. Daarnaast monitort Cito periodiek de normering. Jaarlijks wordt aan de hand van representatieve afnamedata nagegaan of er systematisch verschuivingen in het prestatieniveau plaatsvinden. Indien nodig wordt de normering aangepast.

Conclusie:

Op aspect N1.2.1 wordt aan de toetsen Spelling 3.0 groep 6 het oordeel '**voldoende**' toegekend.

Betrouwbaarheid

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

Bevindingen:

In hoofdstuk 5 van de wetenschappelijke verantwoording wordt beschreven hoe, gebruikmakend van het feit dat de items uit de toets succesvol geschaald zijn met OPLM, een betrouwbaarheidscoëfficiënt, de MAcc ('Accuracy of Measurement'), berekend kan worden die qua interpretatie grote overeenkomst vertoont met de betrouwbaarheidscoëfficiënt uit de klassieke testtheorie. Deze coëfficiënt wordt in de psychometrische literatuur beschreven en als correct aangemerkt.

Conclusie:

De betrouwbaarheidsgegevens worden correct berekend. Op aspect B1.1 wordt aan de toetsen Spelling 3.0 groep 6 het oordeel '**voldoende**' toegekend.

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

Bevindingen:

Voor E5M6, M6, M6E6 en E6 zijn drie betrouwbaarheidsgegevens berekend: standaardmeetfout, MAcc en een gesimuleerde test-hertest betrouwbaarheidscoëfficiënt. Voor E5M6 zijn de gegevens gelijk aan 2,615, 0,90 en 0,90, voor M6 gelijk aan 2,768, 0,92 en 0,92, voor M6E6 gelijk aan 2,639, 0,90 en 0,90 en voor E6 gelijk aan 2,917, 0,91 en 0,91. De auteurs van de wetenschappelijke verantwoording verwijzen naar het beoordelingssysteem van de COTAN waar voor tests die geen zware consequenties voor leerlingen hebben, zoals de toetsen Spelling 3.0 groep 6, een betrouwbaarheidscoëfficiënt van meer dan 0,80 als 'goed' aangemerkt wordt.

Naast klassieke betrouwbaarheidscoëfficiënten is ook de lokale betrouwbaarheid en de meetnauwkeurigheid onderzocht. De betekenis van de meetnauwkeurigheid voor de beslissingen die met de toets genomen worden, is af te leiden uit de betrouwbaarheidstabellen van twee niveauperdelingen (I t/m V en A t/m E) voor E5M6, M6, M6E6 en E6. Tabel 5.2a en tabel 5.2b van de wetenschappelijke verantwoording laten zien dat de leerlingen met de hoogste en de laagste niveau-indicatie voldoende betrouwbaar

geclassificeerd worden. Voor de andere niveaus geldt dat de leerlingen onvoldoende betrouwbaar geclassificeerd worden. Aan de betrouwbaarheidstabellen moet de eis gesteld worden dat alle waarden op de hoofddiagonaal minstens 0,70 zijn omdat de classificaties dan voldoende betrouwbaar zijn in het geval de uitslag op de toets geacht wordt geen belangrijke consequenties voor de leerling te hebben. Heeft de toets wel belangrijke consequenties voor de leerling dan moeten voornoemde waarden minstens 0,80 zijn.

De informatie uit de betrouwbaarheidstabellen kan worden samengevat met twee indices: de plus/minus 1 niveau-index en de marginal classification accuracy. Uit de indices blijkt dat gegeven het gebruiksdoel van de toetsen de classificaties voldoende betrouwbaar zijn, maar dat er wel rekening mee moet worden gehouden dat er altijd sprake zal zijn van misclassificaties ter grootte van maximaal 1 niveau.

Conclusie:

De betrouwbaarheid van de toetsen Spelling 3.0 groep 6 is 'voldoende' als aangenomen mag worden dat de toets geen zware consequenties voor de leerlingen heeft. Op basis van het voorgaande wordt op aspect B1.2 aan de toetsen Spelling 3.0 groep 6 het oordeel '**voldoende**' toegekend.

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 leveren met de toets een uitwerking voor groep 6 van het PO en SO van *Spelling van niet-werkwoorden*, onderdeel van de schriftelijke taalvaardigheid. De werkwoordspelling komt in groep 6 nog te weinig structureel aan bod om al opgenomen te worden in de LVS-toetsing. De toetsen kunnen in groep 6 halverwege of aan het einde van het schooljaar worden afgenomen. Daarnaast zijn voor toetsing 'op maat' en voor leerlingen met een ontwikkelingsachterstand of extra onderwijsbehoeften twee additionele toetsen beschikbaar voor de periode tussen E5 en M6, en tussen M6 en E6.

De toelichting op de inhoud van elk van de toetsonderdelen is uitgebreid en helder.

De moeilijkheidsgraad is, gezien het niveau van de doelgroep, aanvaardbaar. Het antwoordmodel laat geen ruimte voor interpretatie.

De opgavenboekjes geven slechts in één enkel geval aanleiding tot commentaar op itemniveau:

Toets M6E6 Taak 2, item 14 '...cello (spreek uit: *sello*)': dit is het enige voorkomen van een uitspraak-instructie voor de leerkracht in al het toetsmateriaal. Hier had gezocht kunnen worden naar een alternatief doelwoord voor de c-uitgesproken-als-s (als *cel*, *citroen* of *circus*).

De teksten die in de Wetenschappelijke verantwoording opgenomen zijn met betrekking tot de inhoud van de toets, geven een compleet beeld, zijn gedegen en verhelderend.

In de Handleiding voor leerkrachten die bij het toetspakket geleverd wordt, is een heldere inhoudsverantwoording opgenomen. De informatie in dit stuk is conform de informatie die ook in de wetenschappelijke verantwoording is opgenomen. De tekst maakt duidelijk dat het Referentiekader Taal (Domein 4: *Begrippenlijst en Taalverzorging*) en de daarop aansluitende SLO publicatie die de leerstoflijnen spelling beschrijft, in combinatie met een analyse van de meest gebruikte spellingmethoden de bron van de indeling, opbouw en inhoud van de toetsen zijn geweest. Tevens wordt verduidelijkt dat het om indicaties van zowel niveau als progressie van de leerlingvaardigheden gaat. Daarnaast biedt het materiaal voor de leerkracht de mogelijkheid tot, en een handleiding bij, het uitvoeren van een foutenanalyse.

De beoordeling van validiteit beperkt zich hierna tot het statistisch/psychometrisch onderzoek dat verricht is inzake validiteit en dan specifiek begripsvaliditeit.

In de wetenschappelijke verantwoording wordt verslag gedaan van onderzoek naar verschillende aspecten van begripsvaliditeit: unidimensionaliteit, itemkwaliteit, itembias, convergente en divergente validiteit en verschillen tussen relevante subgroepen.

Uit het kalibratieonderzoek van de toetsen blijkt dat het aannemelijk is dat er sprake is van unidimensionaliteit en dat de items één latente trek of construct meten die we, gezien ook de uitkomsten van het hierna genoemde onderzoek, kunnen aanduiden als 'spellingvaardigheid'. Hiermee is voldaan aan de noodzakelijke voorwaarde voor begripsvaliditeit.

Uit het onderzoek naar itemkwaliteit bleek dat de moeilijkheidsgraad van de items op het vooraf gewenste niveau lag, namelijk tussen 0,46 en 0,92. De gemiddelde Rit-waarde is voor alle toetsen 0,38 of hoger waarbij een Rit-waarde van hoger dan 0,30 door de Cotan als goed wordt gekwalificeerd. De kalibratieresultaten op itemniveau en de analyses met betrekking tot de nauwkeurigheid van de itemparameterschattingen bevestigen voorgaande positieve bevindingen.

Onderzoek naar de samenhang tussen de toetsen Spelling met andere LVS toetsen zoals Rekenen-Wiskunde M6, Begrijpend lezen M6, Technisch lezen M6 en Woordenschat M6 liet correlaties zoals verwacht zien, d.w.z. de convergente en divergente validiteit was voldoende. De resultaten geven aan dat er gemeten wordt wat men met de toetsen Spelling 3.0 groep 6 beoogt te meten, namelijk spellingvaardigheid.

Onderzoek naar Differentieel Item Functioneren (DIF) met betrekking tot jongens en meisjes liet zien dat er nauwelijks sprake was van DIF naar sekse. Slechts bij één item bij de toetsen M6 en M6E6 was er sprake van differentieel functioneren met betrekking tot sekse.

Uit onderzoek naar verschillen tussen relevante subgroepen bleken de resultaten aan te sluiten bij de verwachtingen die op grond van theoretische inzichten en eerder onderzoek konden worden geformuleerd: jongere/versnelde leerlingen doen het iets beter dan reguliere leerlingen en die doen het weer iets beter dan oudere/vertraagde leerlingen; meisjes scoren iets hoger dan jongens; leerlingen met een leerlinggewicht van 0,3 en 1,2 scoren op M6 en E6 lager dan leerlingen met leerlinggewicht 0.

Conclusie:

Op aspect V1.1 wordt aan de toetsen Spelling 3.0 groep 6 op dit aspect het oordeel 'voldoende' 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:

Uit de resultaten van het kalibratieonderzoek blijkt dat de items van de verschillende toetsen Spelling 3.0 groep 6 op een eendimensionale vaardigheidsschaal afgebeeld kunnen worden en dat aan de hand van de door de leerling behaalde vaardigheidsscores op de onderscheiden toetsen diens groei adequaat gemeten kan worden.

Conclusie:

Op aspect VA1.1 wordt aan de toetsen Spelling 3.0 groep 6 het oordeel 'voldoende' 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 de Handleiding, met name in hoofdstuk 4, wordt aan de hand van een leerlingrapport de interpretatie van groei op een duidelijke manier beschreven. De 67 procent betrouwbaarheidsintervallen van de vaardigheidsscores worden zowel op het leerlingrapport als in afzonderlijke tabellen vermeld.

In de Handleiding wordt zowel met vergelijkingen tussen verschillende leerlingen als met vergelijkingen op andere tijdstippen van dezelfde leerlingen het volgen van groei van leerlingen adequaat toegelicht.

Conclusie:

Op aspect VA1.2 wordt aan de toetsen Spelling 3.0 groep 6 het oordeel 'voldoende' 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:

Via de portal van Cito B.V. kan gebruik worden gemaakt van rapportageformulieren voor een leerlingrapport, groepsrapport, groepsoverzicht (overzicht van één groep leerlingen tijdens hun schoolperiode) en een alternatief leerlingrapport (voor leerlingen die op een eigen niveau werken).

In de Handleiding wordt in hoofdstuk 7 aandacht besteed aan hoe met ouders over de toetsresultaten gecommuniceerd kan/moet worden. In het bijzonder wordt daarbij gewezen op het leerlingrapport waarin zowel het niveau van de leerling als de progressie

van de leerling numeriek en grafisch gepresenteerd worden. Hiervoor is ook een folder ouderinformatie beschikbaar die men via de website van het Cito kan downloaden. Daarnaast wordt de docent gewezen op misverstanden die zich bij de interpretatie van de niveau-indelingen bij de ouders kunnen voordoen. Ook moeten zij aan ouders het verschil tussen methode-onafhankelijke en methodegebonden toetsen duidelijk maken en erop wijzen dat deze toetsen leerlingen anders (kunnen) beoordelen.

In hoofdstuk 8 worden ook veel gestelde vragen behandeld die weliswaar voor de docenten bestemd zijn maar zij kunnen met die informatie bijvoorbeeld via tienminutengesprekken ook de ouders beter voorlichten.

Conclusie:

Op aspect I1.1 wordt aan de toetsen Spelling 3.0 groep 6 het oordeel '**voldoende**' toegekend.

3. Verzamelstaat

Kwaliteitsaspect	Code	Oordeel
De kwaliteit van de steekproef	<i>S1.1</i>	Voldoende
	<i>S1.2</i>	Voldoende
Normering	<i>N1.1</i>	Voldoende
	<i>N1.2</i>	Voldoende
Betrouwbaarheid	<i>B1.1</i>	Voldoende
	<i>B1.2</i>	Voldoende
Validiteit	<i>V1</i>	Voldoende
Volg-aspect	<i>VA1.1</i>	Voldoende
	<i>VA1.2</i>	Voldoende
Inzicht in leervorderingen	<i>I1</i>	Voldoende

4. Literatuurlijst

- Tomesen, M., Wouda, J. & Horsels, L. (2017). *Wetenschappelijke verantwoording van de LVS-toetsen Spelling 3.0 groep 6*. Arnhem: Cito B.V.
- Cito (2016). *Cito Volgsysteem primair en speciaal onderwijs. Spelling 3.0 groep 6*. Arnhem: Cito