

## 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) Hoofdrekenen 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 Hoofdrekenen is een tempotest en kent twee versies: versie A en versie B.

### Meetpretentie

De SVT Hoofdrekenen meet de vaardigheid in het hoofdrekenen van leerlingen, zoals deze wordt aangeleerd in het Nederlandse basisonderwijs. Daarbij gaat het zowel om de nauwkeurigheid als het tempo van het hoofdrekenen.

### Doelgroep

De SVT Hoofdrekenen is bestemd voor leerlingen in groep 3 tot en met 8 van het Nederlandse basisonderwijs.

### Gebruiksdoel en functie

Het gebruiksdoel van de SVT Hoofdrekenen is vierledig:

- 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 tevens een rol spelen bij het tijdig signaleren van rekenzwakke leerlingen in de eerste schooljaren, waarin het hoofdrekenen in het bijzonder aandacht krijgt.
- De toets kan ten slotte dienen als indicator van een minimumvereiste van 'functionele hoofdrekenvaardigheid'.

### Inhoudelijke theoretische inkadering:

Bij het samenstellen van de rekensommen is er geput uit verschillende rekenmethodes, waarbij gekozen is voor de volgende categorieën: optellen aftrekken, vermenigvuldigen en delen. Bij de keuzes rondom volgorde is er gekozen zich te baseren op de Bredase Rekentoetsen.

### Inhoud van het toetspakket

Het toetspakket bestaat uit:

- Handleiding en Verantwoording
- Instructiekaart
- Rekenblad versie A
- Rekenblad versie B
- Scoremal versie A
- Scoremal versie B
- Informatieblad voor ouders

- Voorbeeld van een Individueel Rapport (3x)
- Voorbeeld van een Individueel Overzicht (2x)
- USB-stick met excelbestand 'Normeringsmodel SVT Hoofdrekenen 2018'. Hierin staan de gegevens over het normeringsmodel van de toets voor de nieuwe normen.

## 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 Pauly K. Berding-Oldersma MSc (secretaris).*

### **De kwaliteit van de steekproef**

#### S1.1. Is de steekproef representatief?

##### *Bevindingen:*

Informatie over de normeringssteekproef wordt in hoofdstuk 6 van de handleiding gepresenteerd. In tabel 6.1 staat dat de normeringssteekproef in totaal 20 scholen, 164 klassen en 3403 leerlingen telt. De representativiteit van de steekproef is op schoolniveau onderzocht met betrekking tot schoolgrootte, regionale spreiding, stedelijkheid, statuscode postcode, denominatie en gemiddelde eindtoetsscore. De representativiteit van de steekproef is op leerlingniveau onderzocht met betrekking tot sekse, leeftijd, achtergrond en leerlinggewichten. De representativiteit voor zowel de variabelen op school- als leerlingniveau is beoordeeld door middel van Cramers V, een maat van samenhang tussen variabelen op nominaal niveau. In plaats van Cramers V had ook Phi als effectmaat (met waarden  $\Phi \leq 0.1$  klein effect volgens Cohen) gerapporteerd kunnen worden op basis van de uitgerekenende chi-kwadraat toetsingsgrootte

Omdat er bij de verdelingen van de didactische leeftijden enige afwijkingen waren ten opzichte van de verdelingen in de populatie voor de regio en achtergrond van de leerlingen, is voor deze variabelen gewogen. Tabel 6.17 geeft de populatie en de gewogen normeringssteekproef voor de variabele regio en Tabel 6.18 geeft de populatie en de gewogen normeringssteekproef voor de variabele achtergrond van de leerlingen. Op deze manier kan namelijk een indicatie worden verkregen in hoeverre de regionale spreiding per leeftijdsgroep en achtergrond van de leerlingen per leerlinggewicht afwijken van de landelijke verdeling en zodoende een motivering worden geven om een gewogen normeringssteekproef te gebruiken voor deze beide variabelen in Tabel 6.17 en Tabel 6.18.

In paragraaf 6.7.2 wordt terecht aandacht besteed aan het feit dat de steekproef gestratificeerd is, waardoor er in feite sprake is van een multilevel probleem in de vorm van schoolafhankelijkheid en er mogelijke substantiële design effecten kunnen ontstaan. Om deze reden werden er in Tabel 6.21 ICC's (Intra Klasse Correlaties) berekend per DL (Didactische Leeftijd) en over de totale normeringssteekproef. Over de totale normeringssteekproef bleek er geen sprake van schoolafhankelijkheid ( $ICC < .05$ ). Voor de verschillende normeringsmomenten bleken de ICC's weliswaar hoger, maar nog steeds laag. Een uitzondering vormde de ICC voor groep 7 (DL = 48) waarvoor een gematigd verband werd gevonden ( $ICC = .223$ ). Bij het opstellen van het continue normeringsmodel werd gecontroleerd voor een eventueel effect van de school en de

groep. Daartoe werd op de gegevens van de normeringssteekproef het normeringsmodel toegepast, zoals beschreven in paragraaf 6.7.1, maar uitgebreid met de variabelen van basisschool, interactie basisschool en normeringsmoment (DL) en interactie basisschool en vaardigheidsscore om te bekijken of deze variabelen een toegevoegde waarde hebben bij de voorspelling van de z-score. Dit bleek echter niet het geval te zijn, waaruit de conclusie kan worden getrokken dat schoolafhankelijkheid geen relevante factor is bij het opstellen van het normeringsmodel.

*Conclusie:*

Op aspect S1.1 wordt aan de Schoolvaardigheidstoets (SVT) Hoofdrekenen groep 3 - 8 het oordeel '**voldoende**' toegekend.

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

*Bevindingen:*

Bij de SVT Hoofdrekenen is er sprake van twee parallelle toetsen, versie A en B. Op pagina 55 van de handleiding staat dat de twee versies binnen een periode van twee weken zijn afgenomen en dat de volgorde van afname gebalanceerd (single group balanced design) is om eventuele volgorde-effecten tegen te gaan. In tabel 5.1 (het 2<sup>e</sup> getal in de onderste rij moet overigens 34.03 zijn i.p.v. 3403) staan de aantallen leerlingen uit de groepen 3 - 8 waarbij de twee versies afgenomen zijn. Groep 5 bestaat uit het grootste aantal leerlingen, namelijk 591, en groep 8 uit het kleinste aantal leerlingen, namelijk 536. In totaal werden de twee toetsen bij 3403 leerlingen afgenomen.

De twee versies meten dezelfde vaardigheid maar uit tabel 5.1 blijkt dat bij de meeste groepen versie A iets moeilijker is dan versie B. Over de totale normgroep bleek het om een verschil van nog geen punt te gaan. Echter met name voor groep 7 was dit verschil groter en bedroeg het 2.5 punten en Cohen's d was bijna 0.1. Via equivalering met de equipercentiel methode met loglineaire presmoothing werd voor de verschillen gecorrigeerd. De resultaten van de equivalering van de toetsscores worden besproken en vermeld in paragraaf 5.3. In Tabel 5.5 worden de gemiddelden en standaarddeviaties gegeven na equivalering met presmoothing, uitgesplitst naar de verschillende normeringsmomenten. Over het algemeen zijn de verschillen klein en moeten verschillen na equivalering worden toegeschreven aan steekproeffluctuaties. De auteur gaat er hierom dan ook vanuit dat versie A en versie B na equivalering als psychometrisch evenwaardig kunnen worden beschouwd.

In Tabel 5.2 op pagina 56 van de handleiding worden de correlaties tussen versie A en versie B gepresenteerd per normeringsgroep en blijkt over de totale steekproef afgerond .97 te zijn. Ook op de verschillende normeringsmomenten blijken de correlaties hoog te zijn, met uitzondering van groep 3. In groep 3 valt de correlatie met .87 relatief laag uit, hetgeen te verklaren is uit het feit dat deze groep minder stabiel is en de beginselen van het hoofdrekenen juist onder de knie krijgen.

Hierbij dient echter wel de kanttekening geplaatst te worden dat de gevonden hoge correlaties in Tabel 5.2 tussen versie A en versie B alleen aangeven dat: (i) op basis van de scores op versie A zeer nauwkeurige voorspellingen gedaan kunnen worden over de scores op versie B en andersom; (ii) de rangordening van leerlingen op grond van de toetsscores van versie A nagenoeg identiek is aan die van versie B. Dit geeft echter geen garantie dat de twee versies als parallel of gelijkwaardig beschouwd mogen worden (een perfecte correlatie hoeft niet noodzakelijkerwijs te betekenen dat de behaalde scores op de twee versies identiek zijn!). Echter in combinatie met de bevinding uit Tabel 5.1 dat zowel de gemiddelden als de standaarddeviaties van beide versies zeer dicht bij elkaar

liggen, is het wel erg aannemelijk om te veronderstellen dat de beide versies als parallel kunnen worden beschouwd.

*Conclusie:*

Het cross-sectionele dataverzamelingsdesign dat hier toegepast is, is adequaat. Op aspect S1.2 wordt aan de SVT Hoofdrekenen groep 3 - 8 het oordeel '**voldoende**' toegekend.

## **Normering**

### N1.2.1. Zijn de normgroepen groot genoeg?

*Bevindingen:*

In tabel 5.1 staan voor de twee versies de aantallen leerlingen van groep 3 – 8 vermeld. De 6 groepen tellen in totaal 3403 leerlingen met een range van 536 – 591 leerlingen.

In paragraaf 6.7 van de handleiding wordt gewezen op de wenselijkheid om de toetsen ook buiten de normeringsmaand te gebruiken. Door middel van continue normering heeft men normen voor de tien verschillende maanden per leerjaar opgesteld. In grafiek 6.1 wordt de continue normering van de SVT Hoofdrekenen grafisch weergegeven.

Er zijn normen geconstrueerd voor de vaardigheidsscores en de accuratessescores. De normen voor de vaardigheidsscores betreffen genormaliseerde standaardcores met een gemiddelde gelijk aan 5 en een standaarddeviatie van 2 (zogenaamde C-schaal). Voor de accuratessescores is gekozen voor stanines waarbij de laatste vier stanines samengenomen zijn vanwege de geringe scoreverschillen tussen de hogere leerjaren, waarbij een grote groep leerlingen nauwelijks tot geen fouten maakt bij hoofdrekenen.

*Conclusie:*

Op aspect N1.2.1 wordt aan de SVT Hoofdrekenen groep 3 - 8 het oordeel '**voldoende**' toegekend.

### N1.2.2. Zijn de normgroepen representatief?

*Bevindingen:*

De representativiteit werd in S1.1 besproken en daar werd geconstateerd dat de normgroepen representatief waren voor relevante variabelen op school- en leerlingniveau.

In de Handleiding wordt vermeld dat de normen in principe geldig zijn tot april 2032. Wanneer uit een tussentijdse steekproef significante verschuivingen blijken kan dit leiden tot eerdere hernormering van de toets.

*Conclusie:*

Op aspect N.1.2.2 wordt aan de SVT Hoofdrekenen groep 3 - 8 het oordeel '**voldoende**' toegekend.

## **Betrouwbaarheid**

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

*Bevindingen:*

In hoofdstuk 7 van de handleiding worden gegevens over de betrouwbaarheid vermeld. Zowel parallelle als test-hertest betrouwbaarheden zijn berekend (betrouwbaarheid aan de hand van interne consistentie kon niet worden berekend, omdat de SVT Hoofdrekenen een tempotest is). De betrouwbaarheden zijn correct berekend, d.w.z. als zijnde respectievelijk de correlatie tussen versie A en versie B en de correlatie tussen de toets op de eerste afname en de toets op de tweede afname na drie weken. Overigens kan hierbij wel de vraag worden gesteld of het wellicht beter was geweest (indien mogelijk uiteraard) om een langere periode dan drie weken te nemen tussen beide afnames om de test-hertest betrouwbaarheid te schatten. Zoals ook geconstateerd door de auteur blijkt namelijk uit Tabel 7.2 dat er sprake is van een aanzienlijk leereffect of herinneringseffect ('carry-over effect'), met name bij groep 8. De toetsaanbieder wijst gebruikers er op dat met dit leereffect rekening gehouden moet worden bij herhaald afnemen van de toets op korte termijn.

#### *Conclusie:*

Op aspect B.1.1 wordt aan de SVT Hoofdrekenen groep 3 - 8 het oordeel '**voldoende**' toegekend.

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

#### *Bevindingen:*

De parallelle betrouwbaarheden voor de 6 afzonderlijke groepen zijn hoog (met uitzondering van groep 3 is de betrouwbaarheid voor alle groepen  $> .90$ ) en voor de totale groep zeer hoog, namelijk  $.968$  (zie Tabel 7.1). De test-hertest betrouwbaarheid (stabiliteit) over de gehele onderzoeksgroep is bijna gelijk aan die van de parallelle betrouwbaarheid, en is eveneens zeer hoog:  $.961$ . Net als bij de parallelle betrouwbaarheid kan geconstateerd worden dat bij de test-hertest betrouwbaarheid de hoogste betrouwbaarheid is te vinden bij groep 8 (en de laagste betrouwbaarheid bij groep 4).

Wat ontbreekt is een schatting van het percentage misclassificaties van de twee niveau-indelingen die gebruikt worden, namelijk niveau A-E en niveau I - V. Gezien de hoge betrouwbaarheden van de toetsen mag aangenomen worden dat het percentage misclassificaties niet groot zal zijn.

#### *Conclusie:*

Wanneer we uitgaan van het door de COTAN gehanteerde criterium (betrouwbaarheid  $> .80$ ), wordt op basis van parallelliteits- en stabiliteitsgegevens aan de SVT Hoofdrekenen groep 3 - 8 op aspect B1.2 het oordeel '**voldoende**' toegekend.

### **Validiteit**

#### VI. 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 toets)?

#### *Bevindingen (inhoudelijk/vakdidactisch):*

Wanneer een toets elk jaar minimaal 1 keer wordt afgenomen, zullen de meetresultaten over de jaren heen een vergelijkbaar geheel gaan vormen. Verder wordt bij deze toets

een beperkt set van items aangeboden. Dit werkt de betrouwbaarheid in de hand. Over de inhoudelijke/vakdidactische validiteit zijn enkele aantekeningen te maken.

De operationalisering van hoofdrekenen zoals deze door Boom is gekozen is beperkt. De beperkingen die worden opgelegd (alles door elkaar w.b. bewerkingen; geen kladpapiergebruik; geen 'mooie opeenvolgende sommen'; geen mogelijkheden voor handig rekenen; etc.) dragen bij aan een eenzijdig beeld van hoofdrekenen. Een en ander raakt aan de discussie over 'met het hoofd' en/of 'uit het hoofd', en over de didactische opdracht die leerkrachten hebben in het rekencurriculum op het gebied van 'handig en flexibel hoofdrekenen'.

Hiernaast is de gekozen/aangeboden toetspraktijk (zes jaar lang steeds dezelfde toets afnemen) een hardhandige/inefficiënte methode om de vaardigheid hoofdrekenen te duiden. Het vermoeden bestaat dat de toets door deze afnamemethode qua motivatie de sterke leerling omhoogstuwt (die krijgt zelfvertrouwen en groeit, ook in een dergelijke toetssetting) en de zwakke leerling naar beneden duwt (die krijgt door de te hoge lat in feite een afstraffing, doordat de nadruk zo op tempo ligt en op een beperkte invulling van het begrip hoofdrekenen).

Er wordt in de volgorde van items rekening gehouden met de ontwikkeling in de zes schooljaren. Daarbij wordt aangesloten bij de gemiddelde leerling, maar er wordt geen rekening gehouden met de 'bandbreedte' aan rekenvaardigheden in de klas.

*Bevindingen (statistisch/psychometrisch, vooral gericht op begripsvaliditeit):*

In hoofdstuk 8 van de handleiding, getiteld 'Begripsvaliditeit', wordt een onderbouwing gegeven van de begripsvaliditeit van de SVT Hoofdrekenen. Argumenten die worden aangedragen hebben betrekking op de lineaire ontwikkeling van het hoofdrekenen gedurende de gehele basisschool (zie Grafiek 8.1), de verschillen tussen jongens en meisjes (jongens doen het gemiddeld 5 punten beter dan meisjes, zie Tabel 8.1), de verschillen tussen leerlingen met een verschillende achtergrond (opvallend genoeg doen leerlingen met een niet-westerse migratieachtergrond het gemiddeld niet slechter of beter op de toets dan de leerlingen met een Nederlandse achtergrond, zie Tabel 8.2) en de relatie tussen de SVT Hoofdrekenen en de Tempo Test Rekenen (lijkt in veel opzichten op de SVT Hoofdrekenen, maar een belangrijk verschil is dat de verschillende rekenoperaties apart worden getoetst). De correlaties voor verschillende groepen tussen voornoemde twee toetsen zijn groter dan .80 (zoals verwacht, iets lager dan de test-hertest betrouwbaarheden; zie Tabel 8.4) en kunnen aangemerkt worden als een bewijs voor soortgenotenvaliditeit (convergente validiteit). De toetsaanbieder geeft aan dat op termijn een correlatieve onderzoek naar de relatie met de SVT Rekenen-Wiskunde zal worden uitgevoerd.

*Conclusie:*

Op aspect V1 wordt aan de SVT Hoofdrekenen groep 3 - 8 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:*

Onderscheiden worden, zie hoofdstuk 4 van de handleiding, de toetsscore (= ruwe score = aantal correct gemaakte sommen) en de accuratessescore (= percentage correct gemaakte sommen). De toetsscore is voor de interpretatie leidend.

In grafiek 4.1 wordt op basis van empirische gegevens, verkregen met behulp van het eerdere vernoemde cross-sectionele dataverzamelingsdesign, het ontwikkelingsverloop aangegeven van de toetsscore en de accuratessescore. Uit de grafiek blijkt dat de gemiddelde toetsscore zich bijna lineair ontwikkelt en dat de accuratessescore na een aanvankelijk sterke stijging in de hogere leerjaren afvlakt.

Elke toetsscore die op een bepaalde toets wordt behaald, correspondeert met een vaardigheidsscore die op een vaardigheidsschaal ligt met een gemiddelde van 5 en een standaarddeviatie van 2. De reden om voor de C-schaal te kiezen is dat andere toetsen van Boom ook deze schaal gebruiken.

*Conclusie:*

Op aspect VA1.1 wordt aan de SVT Hoofdrekenen groep 3 - 8 het oordeel '**voldoende**' toegekend.

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

*Bevindingen:*

Informatie over groei en de interpretatie van groei staat in de paragrafen 4.3 en 4.4 plus bijbehorende losbladige informatie.

Paragraaf 4.3 betreft het individueel rapport waarin van de leerling (numeriek en grafisch) vermeld staan: Score, DLE, Accuratessescore, Accuratesse norm, Percentiel, Niveau I - V, niveau A - E, Vaardigheidsscore met 90% betrouwbaarheidsinterval.

Paragraaf 4.4 betreft het individueel overzicht. In dit overzicht staat van de leerling: Versie van de toets, Afnamedata, DL, Score, Percentiel, DLE, Vaardigheidsscore met 90% betrouwbaarheidsinterval. Hierbij kan nog opgemerkt worden dat het gebruik van vaardigheidsscores meer voor de hand liggen dan DLE's, omdat met grotere precisie de ontwikkeling van de leerling kan worden gevolgd. Maar omdat er nog steeds vraag is naar DLE's, worden deze ook beschikbaar gesteld. Het wordt door de auteur overgelaten aan de verantwoordelijkheid van de gebruiker om hierin een eigen afweging te maken. Daarnaast bevat het overzicht van de SVT Hoofdrekenen een grafiek met op de x-as de DL en op de y-as de vaardigheidsscore met bijbehorend betrouwbaarheidsinterval.

*Conclusie:*

Op aspect VA1.2 wordt aan de SVT Hoofdrekenen groep 3 - 8 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:*

In de handleiding zijn vijf casussen opgenomen waarin individuele rapporten en individuele overzichten besproken worden. Historische overzichten kunnen bij het Boom testcentrum opgevraagd worden. Hoe deze overzichten gecommuniceerd worden met

ouders en verzorgenden wordt niet vermeld. Bij het testcentrum kunnen ook groepsrapporten, vergelijkbaar met het individuele rapport, aangemaakt worden.

N.B.: De schaal van de percentielscore op pagina 42, 44 en 46 is niet erg gelukkig weergegeven. Het lijkt nu net of de schaal van de normaal verdeling die er boven staat dezelfde is als de schaal met het label "Percentiel" die er direct onder staat. Echter, het probleem met percentielscores is dat de verdeling van de leerlingen over de scoreschaal vaak niet gelijkmatig is. Meestal clusteren de scores van de meeste leerlingen zich in de buurt van het 50e percentiel. Percentielen komen dicht op elkaar voor, rond het midden en liggen verder uit elkaar aan het begin en einde. Gelijke verschillen op de ruwe scoreschaal, zoals bijvoorbeeld: 50-60 en 80-90, leiden daardoor niet tot gelijke verschillen op de percentielschaal. Het verschil van 10 punten bij 50-60 kan resulteren in een verschil van enkele tientallen punten op de percentielschaal, terwijl het verschil van 10 punten bij 80-90 maar enkele percentiepunten zal bedragen. Op basis van een verschil in percentielscores kan hooguit geconstateerd worden welke leerling beter is, maar niet hoeveel beter. Een percentiel is dus niet gebaseerd op een intervallschaal, maar geeft alleen een rangorde aan en is dus gebaseerd op een ordinale schaal. Om verwarring bij de gebruiker te voorkomen verdient het aanbeveling om in de toekomst bij volgende versies van de handleiding expliciet te vermelden dat de twee schalen niet een-op-een op elkaar zijn afgebeeld.

Er is een smalle definitie van hoofdrekenen gekozen voor deze toets. De beperkingen die worden opgelegd (alles door elkaar w.b. bewerkingen; geen kladpapier-gebruik; geen 'mooie opeenvolgende sommen'; geen mogelijkheden voor handig rekenen; etc.) dragen bij aan een eenzijdig beeld van hoofdrekenen. Dit kan tot gevolg hebben dat er beeldvorming kan ontstaan (richting ouders, e.a.) als de beoordeling van de kwaliteit van hoofdrekenen van kinderen zich zou beperken tot de Boom SVT-hoofdrekenen. Dan kan de misvatting ontstaan dat de score op de SVT-hoofdrekenen de actuele stand van de hoofdrekenvaardigheid bij dat kind representeert.

*Conclusie:*

Op aspect I.1 wordt aan de SVT Hoofdrekenen groep 3 - 8 het oordeel '**voldoende**' toegekend.

### 3. Verzamelstaat

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

### 4. Literatuurlijst



Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.), p. 222, table 7.2.3.

Vos de, T. (2018). *Handleiding en Verantwoording Schoolvaardigheidstoets Hoofdrekenen*. Amsterdam: Teije de Vos & Boom uitgevers.