



Aansluiten bij de leerling

Implementatie van de ADIT in het eerste
leerjaar van het voortgezet onderwijs

Bachelor Afstudeerscriptie

Kyra van Rinsum
Studentnummer: 325879
325879@student.saxion.nl

*Toegepaste psychologie
Academie Mens en Arbeid*

Arnhem, 13 – 03 – 2017



**Onderzoeksrapport:**

Adaptief testen: Implementatie van de ADIT in het eerste leerjaar van het voortgezet onderwijs.

Over de auteur:

Naam: Kyra van Rinsum

Studentnummer: 325879

Klas: DTP3VF

E-mailadres: 325879@student.saxion.nl

Hogeschool Saxion, Academie Mens en Arbeid, Toegepaste Psychologie

Opdrachtgevers:

A-VISION Advisering en Begeleiding B.V.

Mevr. A. Lubbers, Directrice A-VISION

Afstudeerbegeleider:

1^e begeleider: Dhr. G. Roemer
Hogeschool Saxion,
Academie Mens en Arbeid, opleiding Toegepaste Psychologie

2^e begeleider: Dhr. J. Veldhuijzen
Hogeschool Saxion,
Academie Mens en Arbeid, opleiding Toegepaste Psychologie

Angelique Egberink
A-VISION Advisering en Begeleiding B.V.
Methodoloog

13 maart 2017

Voorwoord

Voor u ligt de scriptie “Aansluiting bij de leerling: Implementatie van de ADIT in het eerste leerjaar van het voortgezet onderwijs”. Deze scriptie is in het kader van de afronding van de studie Toegepaste Psychologie (TP) aan Saxion te Deventer geschreven en is tot stand gekomen in opdracht van onderwijsadviesbureau A-VISION.

Tijdens mijn opleiding had ik al duidelijk waar mijn interesse lag: de (online) diagnostiek. Ik wist dan ook zeker dat ik mijn afstudeeropdracht in de diagnostiek wou doen. Tijdens mijn stage kwam ik gelukkig A-VISION tegen en na een mail werd ik uitgenodigd voor een oriënterend gesprek. Sindsdien is het snel gegaan. Voordat ik begon aan deze opdracht had ik echter nog geen ervaring met of kennis over digitaal testen. Laat staan adaptief testen. Maar juist omdat ik hier nog geen ervaring mee had, kon ik mij verdiepen in mijn rol als onderzoeker. Dit heeft geleid tot een leerzame periode en opdracht. Het heeft mij een compleet beeld opgeleverd van het proces van online en adaptief testen, waarnaar onderzoek is verricht.

Graag zou ik hierbij de mogelijkheid willen nemen om de volgende mensen te bedanken voor hun bijdrage aan dit onderzoek. Allereerst wil ik natuurlijk beginnen met het bedanken van de scholen, leerkrachten en leerlingen bij wie ik dit onderzoek uit mocht voeren. Zonder hun was dit onderzoek niet mogelijk geweest. Daarnaast wil ik A-VISION bedanken, die mij de kans heeft geboden om mijn afstudeeropdracht bij hun bedrijf uit te mogen voeren. Zonder de steun, het meedenken en de samenwerking van en met de medewerkers van A-VISION, was deze scriptie niet tot stand gekomen. Hierbij wil ik in het bijzonder drie mensen bedanken: John, die ondanks zijn drukke schema tijd voor mij vrij heeft gemaakt voor vragen en voor het verzamelen van de benodigde gegevens; Astrid, die toen ik dacht dat dit onderzoek niet haalbaar voor mij was, toch bleef geloven in een goede afloop en met mij mee heeft gedacht tijdens dit proces; en natuurlijk Angelique die mij vanuit A-VISION heeft begeleid gedurende mijn afstudeerperiode, bij wie ik met vragen terecht kon en bij wie ik terecht kon voor ideeën en oplossingen. Daarnaast wil ik ook mijn begeleiders vanuit de opleiding bedanken. Mijn tweede begeleider de heer Veldhuijzen die kritisch naar mijn hoofdstukken wilde kijken wanneer dit nodig was en bij wie ik terecht kon met eventuele vragen. Maar natuurlijk ook mijn eerste begeleider de heer Roemer die mij, tijdens mijn ups-and-downs, heeft begeleid. Die kritische opmerkingen heeft gemaakt ten opzichte van mijn schrijfstijl, de opbouw en uitvoer van mijn scriptie, maar wie ik ook altijd lastig kon vallen met vragen. Tot slot wil ik afsluiten door iedereen thuis te bedanken: mijn moeder, vader, Djoewi, mijn zusjes, Owen en lieve vriendin. Zij hebben mij gedurende mijn scriptietraject geholpen wanneer ik ergens niet direct uit kwam en wanneer ik een steuntje of duwtje in de rug nodig had. Daarnaast zijn zij altijd geduldig met mij geweest wanneer ik gestrest was, of wanneer ik na een lange dag minder vrolijk thuiskwam. Om het helemaal compleet te maken wil ik ook Dobby en Bella graag bedanken voor de liefde en steun die ze me altijd geven. Dank jullie wel.

Ik wens u veel leesplezier.

Kyra van Rinsum

Arnhem, 13 maart 2017.

Samenvatting

Differentiëren en aansluiten bij de leerling wordt in het huidige onderwijs steeds belangrijker. Hier wil A-VISION, onderwijsadvies bureau te Apeldoorn, op inspelen. Dit willen ze realiseren met de, door hun voor groep 8 ontwikkelde, Adaptieve Digitale Intelligentietest (ADIT). Om de ADIT breder inzetbaar te maken, moet er onderzocht worden of de huidige itembank over voldoende items beschikt om het niveau van de nieuwe doelgroep nauwkeurig in te schatten. Dit onderzoek heeft als doelstelling te achterhalen in hoeverre de huidige ADIT te implementeren is in het eerste leerjaar van het VO.

De ADIT is gebaseerd op ontwikkelingen vanuit de Item Response Theory (IRT). De IRT gaat ervan uit dat de meting van een eigenschap van een persoon met een bepaald item de grootste nauwkeurigheid heeft wanneer de moeilijkheid van het item en het niveau van de persoon op de psychologische eigenschap samenvallen. Dit gebeurt bij een subjectieve kans van 0,5. Door alleen items voor te leggen waarbij de subjectieve kans het dichtst bij de 0,5 ligt, ontstaat er een adaptieve test. De ADIT kan deze items aanbieden door, o.a. door middel van de moeilijkheidsparameter (B), het vaardigheidsniveau van de respondent in te schatten.

Om te onderzoeken of de huidige itembank over voldoende items beschikt, werd de ADIT aan 406 havo en vwo-leerlingen uit leerjaar 1 en 2 voorgelegd. Aansluitend hebben deze leerlingen ook een enquête ingevuld over hun ervaring met de ADIT. De hoofdvraag van het onderzoek is beantwoord door het percentage leerlingen te berekenen die de 'moeilijkste' items aangeboden hebben gekregen. Daarnaast zijn in de beantwoording van de hoofdvraag de TIQ-scores en de resultaten uit de enquête meegenomen.

Uit de resultaten blijkt dat de subtesten, op twee subtests na, over onvoldoende items beschikken om het niveau van de leerlingen goed in te kunnen schatten. Hierbij zijn er statistisch significante verschillen gevonden tussen havo en vwo-leerlingen en leerlingen uit leerjaar 1 en leerjaar 2. Uit de enquête is gebleken dat de leerlingen de afname van de ADIT over het algemeen als positief hebben ervaren. Wel gaven ze aan voor het beantwoorden van de vragen te weinig tijd te hebben ervaren.

Uit de resultaten kan geconcludeerd worden dat de huidige ADIT itembank over te weinig items beschikt om het niveau van eerstejaars havo/vwo leerlingen nauwkeurig in te kunnen schatten. De huidige ADIT-itembank zou aangevuld moeten worden met items met een hogere moeilijkheidsgraad, om de ADIT te kunnen implementeren in het eerste leerjaar van het voortgezet onderwijs. Vervolgonderzoek naar het uitbreiden van de itembank zou zich specifiek kunnen richten op vwo-leerlingen, aangezien deze in dit onderzoek minder vertegenwoordigd zijn dan havo-leerlingen.

Inhoudsopgave

Verklarende woordenlijst	Pg. 8
1. Inleiding	Pg. 9
1.1. De aanleiding	Pg. 9
1.2. Centrale onderzoeksvraag en deelvragen	Pg. 9
2. Theoretisch kader	Pg. 11
2.1. Intelligentietesten	Pg. 11
2.2. Adaptief testen door middel van de Item Response Theory	Pg. 11
2.3. Moeilijkheidsparameter	Pg. 11
2.4. De adaptieve testprocedure	Pg. 12
2.5. Waarom adaptief testen?	Pg. 13
2.6. De ADIT	Pg. 13
2.7. Conceptueel model	Pg. 14
3. Methoden van onderzoek	Pg. 15
3.1. Onderzoeksopzet	Pg. 15
3.2. Doelgroep	Pg. 15
3.3. Onderzoeksinstrumenten -ADIT	Pg. 15
3.4. Onderzoeksinstrumenten -Enquête	Pg. 16
3.5. Onderzoeksprocedure	Pg. 16
3.6. Analyses	Pg. 16
3.7. Betrouwbaarheid en validiteit onderzoeksdesign en -instrumenten	Pg. 17
4. Onderzoeksresultaten	Pg. 19
4.1. Respons en uitvoering- ADIT	Pg. 19
4.2. Respons en uitvoering- Enquête	Pg. 19
4.3. Resultaten deelvraag 1	Pg. 20
4.4. Resultaten deelvraag 2	Pg. 22
4.5. Resultaten deelvraag 3	Pg. 23
5. Conclusie, discussie en aanbevelingen	Pg. 24
5.1. In hoeverre bevat de huidige ADIT-itembank voldoende items om een, vooraf door A-VISION bepaalde, nauwkeurige TIQ meting te geven voor leerlingen uit het eerste leerjaar van het VO?	Pg. 24
5.2. In hoeverre bevat de huidige ADIT-itembank voldoende items met een hoge moeilijkheidsgraad om het intelligentieniveau van leerlingen uit het eerste leerjaar VO, nauwkeurig in te schatten?	Pg. 24
5.3. Hoe wordt de afname van de ADIT door de leerlingen ervaren op de punten: lengte van de test, vraagstelling, vormgeving en de ervaren moeilijkheid?	Pg. 25
5.4. Conclusie	Pg. 25
5.5. Discussie	Pg. 26
5.6. Aanbevelingen	Pg. 27
5.6.1. Aanbevelingen voor experimentele basis	Pg. 27
5.6.1. Aanbevelingen voor vervolgstappen	Pg. 28
5.7. Afsluiting	Pg. 28
Literatuurlijst	Pg. 29
Bijlagen	Pg. 32
B1. Overzicht intelligentietesten en rol ADIT	Pg. 32
B2. Biografische informatie populatie	Pg. 33
B3. Standaard e-mail contact scholen	Pg. 34
B4. Testinstructie ADIT	Pg. 35

B5. 'Ben ik geschikt'-stappenplan	Pg. 37
B6. Enquête	Pg. 39
B7. Overzicht afnameprocedure ADIT	Pg. 40
B8. Normaalverdeling TIQ	Pg. 41
B9. Testinformatiecurve	Pg. 42
B10. Uitkomsten enquête	Pg. 44
B11. Dankwoord scholen	Pg. 46

Verklarende woordenlijst

Clustersteekproef: Er is sprake van een clustersteekproef wanneer er een aselechte steekproef uit deelpopulaties ('clusters') van de gehele onderzoekspopulatie wordt getrokken.

COTAN: Staat voor Commissie Testaangelegenheden Nederland en deze commissie beschrijft en beoordeelt testen die in Nederland worden gebruikt.

Discriminerend vermogen: Het discriminerend vermogen van een item geeft aan of dit item onderscheid kan maken tussen laag en hoog scorende respondenten.

Enkelvoudige variantieanalyse ANOVA: Dit is een toetsingsprocedure in SPSS om na te gaan of steekproefgemiddelden van meer dan twee groepen van elkaar variëren.

Getrapte steekproef: Bij een getrapte steekproef wordt er uit een clustersteekproef weer een aselechte groep gehaald. De steekproeftrekkingen worden in verschillende stadia, na elkaar en afzonderlijk verricht.

Non-verbale intelligentie: De non-verbale intelligentie meet hoe iemand praktisch omgaat met zijn of haar kennis, probleemoplossend vermogen en ruimtelijk inzicht zijn hier voorbeelden van.

T-toets: De T-toets is een statistische toets die gebruikt kan worden om na te gaan of er een verschil is tussen twee populatiegemiddelden.

Verbale intelligentie: Verbale intelligentie heeft betrekking op bijvoorbeeld de woordenschat en het taalgevoel van een persoon.

1. Inleiding

In dit hoofdstuk wordt een aanleiding gegeven voor dit onderzoek. Daarnaast worden de daarbij behorende onderzoeksvragen en de doelstelling van dit onderzoek besproken.

1.1. De aanleiding

Om alle kinderen een passende onderwijsplek te geven en om de ontwikkeling van de leerlingen in brede zin te stimuleren, is differentiëren belangrijk (Coubergs, Stuyven, Engels, Cools, & de Martelaer, 2013). In het huidige onderwijs gebeurt dit al, maar het onderwijs van de toekomst wil meer. In het Nationaal Onderwijsakkoord gepubliceerd in september 2013 werd er gesproken over ambities en doelen voor het onderwijs voor de toekomst. Hierin staat onder andere dat in 2020 de talenten van middelbare scholieren optimaal worden benut en dat zij ruimte hebben om te differentiëren in niveau en tempo. Zo zou een havo-leerling die erg goed is in Engels zijn examen bijvoorbeeld in 2020 op vwo-niveau kunnen laten afnemen (Sectorakkoord VO 2014-2017, 2014). Differentiatie wordt dus in de toekomst steeds belangrijker.

Kennis over zowel de cognitieve kenmerken (eerdere prestaties en kennis) als niet-cognitieve kenmerken (motivatie, houding etcetera) van de leerlingen zijn, om goed te kunnen differentiëren, belangrijk (Tomlinson & Imbeau, 2010). Een van de manieren om achter de cognitieve kenmerken te komen is door bij de leerlingen een intelligentietest af te nemen. Of een leerling bijvoorbeeld gebaat is bij extra mondelinge toelichting kan voortkomen uit een verschil tussen het verbale en non-verbale IQ. Zo kan een intelligentietest bijvoorbeeld inzicht geven in een betere aansluiting van de onderwijsvorm op de leerling. De meest bekende intelligentietesten voor middelbare scholieren zijn op dit moment de NIO, GIVO, GIT en de WISC. Ondanks dat de meeste intelligentietesten op papier worden afgenomen, neigen recentere ontwikkelingen meer naar adaptief en via de computer af te nemen intelligentietesten. A-VISION wil inspelen op de huidige ontwikkelingen in het onderwijs en een alternatief bieden voor de huidige gebruikte intelligentietesten voor deze doelgroep door middel van de ADIT. De ADIT (Adaptieve Digitale Intelligentietest) heeft als voordeel dat hij adaptief is en zich aanpast aan het niveau van de leerlingen. Het grootste voordeel van adaptief testen, in vergelijking met niet adaptief toetsen, is de verkorte toetstijd. Doordat een adaptieve toets uitsluitend opgaven bevat die toegesneden zijn op het niveau van de leerling, bevat deze tot 40% minder opgaven dan een niet-adaptieve toets die even nauwkeurig meet. Ook wordt een adaptieve toets als minder frustrerend ervaren door leerlingen omdat ze immers worden uitgedaagd op hun eigen niveau, en niet veel hoger of veel lager (van Alem & ter Horst, 2013; Cito, n.d.). Er is echter ook literatuur te vinden die het tegendeel beweert (Schoonman, 2011). Hier wordt in hoofdstuk 2 verder aandacht aan besteed. Doordat de ADIT digitaal afgenomen kan worden, hoeft de test niet per se door een orthopedagoog of gekwalificeerd psychodiagnosticus afgenomen te worden. Een leerkracht of intern begeleider op school is ook in staat om de ADIT bij de leerling af te nemen op een computer of laptop. Dit geldt echter alleen als de persoon bekend is met de test en bijbehorende testinstructies. Dit is kosten- en tijdbesparend. De ADIT zou om deze redenen een toevoegende waarde kunnen hebben in het middelbaar onderwijs. Adaptief testen heeft echter ook nadelen. Op deze voor- en nadelen wordt in hoofdstuk 2 ingegaan. Op dit moment is de ADIT echter alleen COTAN beoordeeld voor groep 8 leerlingen. De ADIT heeft een goede COTAN-beoordeling op de onderdelen; kwaliteit van de handleiding en betrouwbaarheid. Op de overige onderdelen heeft de ADIT een voldoende COTAN-beoordeling. Om de ADIT breder inzetbaar te maken, moet er onderzocht worden of de huidige itembank over voldoende items beschikt om het niveau van leerlingen uit het eerste leerjaar van het voortgezet onderwijs nauwkeurig te kunnen inschatten.

Dit onderzoek wordt uitgevoerd in opdracht van onderwijsadviesbureau A-VISION. Dit is een bedrijf dat is gespecialiseerd in het ontwikkelen van digitale, adaptieve testen en online testplatforms. Daarnaast verzorgt A-VISION individuele of groepsonderzoeken, advisering en trainingen passend bij de individuele ontwikkelings- en leervraagstukken. Het bedrijf heeft als visie om elk kind op de juiste plek in het onderwijs te krijgen. Door inzicht te creëren in de mogelijkheden van de leerling kan er een schoolkeuze aangeboden worden met de meeste mogelijkheden voor de toekomst. Het motto van het bedrijf luidt daarom ook: A-VISION met het oog op de toekomst (A-VISION, n.d.).

1.2. Centrale onderzoeksvraag en deelvragen

Intellectuele vaardigheden van individuen ontwikkelen zich door tot in de adolescentie. Met het ouder worden zijn individuen dus, over het algemeen, steeds beter in staat moeilijkere en complexere cognitieve problemen op te lossen (Resing & Drenth, 2007). Een test bestaat uit verschillende items die zijn samengevoegd in een itembank. Er zitten voor groep 8-leerlingen voldoende items in de huidige ADIT-itembank om hun niveau nauwkeurig in te schatten. De huidige itembank zou, gebaseerd op de theoretische kennis, te weinig items bevatten om het niveau van leerlingen uit het eerste leerjaar van het VO nauwkeurig in te kunnen schatten. Een mogelijke oplossing hiervoor is het uitbreiden van de itembank met vragen met een hogere moeilijkheidsgraad die beter passen bij het niveau van havo/vwo-leerlingen uit het eerste leerjaar.

Om dit te onderzoeken is er gekozen voor een kwantitatieve onderzoeksmethode waarvoor er 158 scholen zijn benaderd om met hun havo- en vwo-leerlingen van het eerste en tweede leerjaar mee te doen met dit onderzoek. Bij deelnemende leerlingen wordt de ADIT afgenomen. De onderzoekszetel en de keuze voor onderzoeksdoelgroep zijn uitgewerkt in hoofdstuk 3. Met dit onderzoek wordt getracht antwoord te geven op de volgende onderzoeksvraag en deelvragen:

In hoeverre is de huidige ADIT geschikt voor implementatie in het eerste leerjaar van het voortgezet onderwijs?

- Deelvraag 1: In hoeverre bevat de huidige ADIT-itembank voldoende items om een, door A-VISION vooraf bepaalde, nauwkeurige TIQ-meting te geven voor leerlingen uit het eerste leerjaar havo/vwo van het VO?
- Deelvraag 2: In hoeverre bevat de huidige ADIT-itembank voldoende items met een hoge moeilijkheidsgraad ($B \geq 0$) om het intelligentieniveau van leerlingen uit het eerste leerjaar havo/vwo van het VO nauwkeurig in te schatten?
- Deelvraag 3: Hoe wordt de afname van de ADIT door de leerlingen ervaren op de punten: lengte van de test, vraagstelling, vormgeving en de ervaren moeilijkheid?

In dit onderzoek wordt onder implementatie de volgende definitie verstaan, zoals gegeven door Daamen (2015):

Implementatie is een procesmatige en/of planmatige invoering van een vernieuwing of verandering. Met als smal doel dat de vernieuwing of verandering is geïntegreerd in het beroepsmatige handelen, in het functioneren van organisatie(s) of in de structuur van de sector. En als breed doel dat de vernieuwing of verandering verbetering oplevert. (Hoofdstuk 1, para. 2)

Dit onderzoek heeft als doelstelling te achterhalen in hoeverre de huidige ADIT te implementeren is in het eerste leerjaar van het voortgezet onderwijs en het doen van aanbevelingen voor de implementatie. Hierdoor kan onderzocht worden wat de mogelijkheden en onderwijsbehoeften van de leerling zijn. Wanneer dit bekend is, kan hierop worden ingespeeld en kan er beter passend onderwijs aangeboden worden. De uitspraken van dit onderzoek zullen betrekking hebben op havo en vwo-leerlingen uit het eerste leerjaar van het voortgezet onderwijs. De doelgroep van dit onderzoek zijn echter havo en vwo-leerlingen uit het eerste en tweede leerjaar van het voortgezet onderwijs. In de volgende hoofdstukken wordt er een theoretisch kader geschetst voor het onderzoek en wordt de keuze voor deze doelgroep en het onderzoek uitgewerkt.

2. Theoretisch kader

Er zal in dit hoofdstuk worden ingegaan op de constructen die bij dit onderzoek van belang zijn. Allereerst zal er worden ingegaan op de principes en de werking van het adaptief testen. Daarnaast zullen de technische aspecten met betrekking tot de ADIT worden weergegeven. Dit hoofdstuk wordt afgesloten met een conceptueel model.

2.1. Intelligentietesten

Een intelligentietest gebruikt men wanneer de cognitieve capaciteiten van een individu in beeld gebracht moeten worden. Dit is vaak een onderdeel van diagnostisch onderzoek. Op middelbare scholen worden intelligentietesten veelal ingezet voor het vormen van een ontwikkelingsperspectief plan (OPP) of wanneer er sprake is van leervragen. Zo kan een intelligentietest ingezet worden wanneer onduidelijk is waarom een leerling niet goed presteert op zijn huidige niveau. Aangenomen wordt namelijk dat kinderen met hoge intelligentieniveaus schoolkennis zich sneller en makkelijker eigen maken, hogere cijfers behalen en het op school leuker vinden (Resing & Drenth, 2007). Het voorspellen van schoolsucces en -prestaties is echter een complex proces waarin veel meer factoren dan het IQ een rol spelen. Het verband tussen IQ-scores en schooltoetsen ligt gemiddeld rond de .50. Dit betekent dat andere factoren zoals concentratie, motivatie en culturele achtergrond, in combinatie met intelligentie, een betere voorspelling kunnen geven (van Dijk, 2012). Zowel Pameijer (2014) als Peter Tellegen (Valk, 2005) geven aan dat een intelligentietest, mits goed gebruikt, toch een belangrijke bijdrage kan leveren bij onderwijsvraagstukken: hij moet worden afgenomen in het belang van het kind, moet goed ontworpen zijn en de uitkomsten moeten gerelativeerd worden. In de praktijk wordt vaak aangeraden alleen tests te gebruiken die door de COTAN op zeven criteria als goed of voldoende zijn beoordeeld (Resing & Drenth, 2007). Voor het middelbaar onderwijs bestaan er meerdere COTAN beoordeelde intelligentietesten. Hiervan is een overzicht te vinden in bijlage 1 (B1).

2.2. Adaptief testen door middel van de Item Response Theory

Veel intelligentietesten zijn geconstrueerd volgens de principes van de klassieke testtheorie. Bij deze klassieke manier van testen wordt er een, voor eenieder gelijke, test aan de respondenten voorgelegd (Drenth & Sijsma, 2012). Hiermee wordt bedoeld dat iedereen dezelfde items, met dezelfde oplopende moeilijkheid, in dezelfde volgorde voorgelegd krijgt. Sinds het begin van de jaren zestig van de vorige eeuw werd in de testdiagnostiek echter steeds meer de nadruk gelegd op moderne modellen die vaak worden vertegenwoordigd in de Item Response Theory (IRT). Verondersteld wordt dat de IRT een meting met hogere of dezelfde nauwkeurigheid kan geven als de klassieke testmodellen, maar met een kortere afnametijd (Drenth & Sijsma, 2012; Lai, et al., 2011).

Bij intelligentietesten die zijn ontwikkeld volgens de klassieke testtheorie beginnen alle respondenten met hetzelfde startitem en geleidelijk aan worden er items voorgelegd die moeilijker zijn (Drenth & Sijsma, 2012). Dit betekent, vanuit het oogpunt van een zo groot mogelijke nauwkeurigheid van de meting, dat de respondenten hier meestal geen optimaal geschikte items voorgelegd krijgen. Adaptief testen door middel van de IRT houdt hier rekening mee en gaat uit van het volgende uitgangspunt: de meting van een eigenschap van een persoon met een bepaald item heeft de grootste nauwkeurigheid wanneer de moeilijkheid van het item en het niveau van de persoon op de psychologische eigenschap samenvallen. De persoon heeft, met het item dat het best pas bij zijn/haar niveau, een subjectieve kans die gelijk is aan 0,5 ($p=0,5$). (Drenth & Sijsma, 2012). Vanuit de persoon gezien is de kans op een goed antwoord even groot als de kans op een fout antwoord. Het item heeft dus een middelmatige moeilijkheid. Door een respondent alleen die vragen voor te leggen waarbij de subjectieve kans het dichtst bij de 0,5 ligt, ontstaat er een adaptieve test (Lai, et al., 2011).

2.3. Moeilijkheidsparameter

De IRT gaat bij het ontwikkelen van een adaptieve test uit van latente variabelen. Dit zijn onobserveerbare gehelen die de observeerbare variabelen beïnvloeden (Montgomery & Cutler, 2013). Zo kan het rekenvaardigheidsniveau (onobserveerbaar) van een individu de testscore op een intelligentietest (observeerbaar) beïnvloeden. De IRT gaat ervan uit dat wanneer dit waar is, dit andersom ook zou betekenen dat de testscore, of de respons die een persoon op een item geeft, een indicator kan zijn voor het vaardigheidsniveau van een individu (Embretson & Reise, 2000). Bij een adaptieve test wordt daarom, op basis van de antwoorden op en/of de eigenschappen van de items, vaardigheidsniveaus ingeschat. Deze inschatting wordt gedaan met behulp van parameters. Bij de ADIT wordt, voor het inschatten van het vaardigheidsniveau, gebruik gemaakt van twee

parameters: de moeilijkheidsparameter (de parameter die de moeilijkheid van een item weergeeft) en de discriminatieparameter (de parameter die het discriminerend vermogen van een item weergeeft). Er zal tijdens dit onderzoek en bij het analyseren van de resultaten alleen gebruik worden gemaakt van de moeilijkheidsparameter. Deze zal hieronder worden toegelicht.

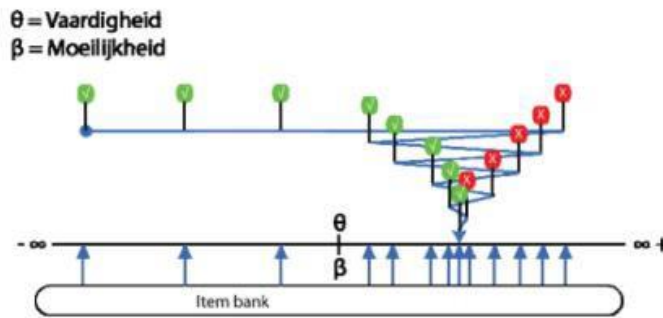
De moeilijkheidsparameter (B) geeft de moeilijkheid van een item weer. Bij A-VISION ligt de moeilijkheidsparameter ongeveer op een schaal van -3 tot 3 ($-3 \geq B \leq 3$). Of een item als moeilijk ervaren wordt, verschilt per persoon. De leeftijd en het niveau van een persoon zijn factoren die meebepalen of een item een p-waarde krijgt van 0,5. Moeilijkheidsparameters worden om deze reden van tevoren vastgesteld met behulp van de items in een normeringsonderzoek die aan de populatie is voorgelegd (Montgomery & Cutler, 2013; Carson, Gillon & Boustead, 2013). Hieruit ontstaat er een schaal waarop de items op hun moeilijkheid voor de populatie worden weergegeven. Deze schaal is een constante voor deze populatie (Montgomery & Cutler, 2013).

2.4. De adaptieve testprocedure

De ADIT maakt gebruik van een gestructureerde testafname. De subtesten worden dus in een vaste volgorde afgenomen. De respondenten krijgen aan het begin, per subtest, een item gepresenteerd dat in de populatie een gemiddelde moeilijkheid ($\delta = 0$) heeft. Er kan pas een schatting van het vaardigheidsniveau worden gemaakt wanneer er minimaal 1 item goed en 1 item fout is beantwoord. Wanneer hieraan is voldaan maakt de computer op basis van de itemscores een eerste, nog erg onnauwkeurige, schatting van het vaardigheidsniveau van de persoon (Lai, et al., 2011). Geeft de respondent alleen maar foute antwoorden, of alleen maar goede antwoorden, dan wordt het vaardigheidsniveau dat wordt gebruikt bij het bepalen van het volgende item telkens met 0,5 verhoogd dan wel verlaagd, totdat een goed én een fout antwoord gegeven zijn (Vegt & Metselaar, 2013). Op ieder punt van de latente schaal is het mogelijk om de informatiewaarde van elk item te berekenen. De items met de hoogste informatiewaarde resulteren tot de grootste reductie in standaardmeetfout van de schatting en hebben dus een grotere betrouwbaarheid. Omdat de schatting afhankelijk is van het goed of fout beantwoorden van items zijn de informatiewaarden van de items voor iedere respondent anders. Het item met de grootste informatiewaarde wordt gekozen als volgend item zodat de moeilijkheid het beste past bij het geschatte vaardigheidsniveau. Dit resulteert weer in een nieuwe schatting van het vaardigheidsniveau die wordt gebaseerd op de scores van het startitem en het nieuwe item (Weiss & Kingsbury, 1984; Carson, Gillon & Boustead, 2013). Doordat de test inmiddels een extra item bevat gaat hierdoor de nauwkeurigheid van de schatting omhoog. Over het algemeen geldt immers dat een groter aantal items leidt tot een grotere nauwkeurigheid van de meting (Drenth & Sijsma, 2012). Op basis van deze schatting wordt weer een volgend item uit de itembank gekozen. Dit gaat zo door totdat een vaardigheidsniveau is geschat die een, van tevoren met de Maximum Likelihood schatter vastgestelde en bevredigende, nauwkeurigheid heeft. Het aanbieden van items stopt dan. Het gaat buiten het bereik van deze scriptie om, om de Maximum Likelihood schatter hier te bespreken. Omdat er per subtest informatie wordt verzameld, zijn er voor de ADIT na ieder item een aantal voorwaarden. Op basis van de voorwaarden wordt besloten om de subtest af te sluiten dan wel voort te zetten. (Vegt & Metselaar, 2013). De voorwaarden zijn:

- Er wordt altijd een volgend item aangeboden zolang er minder dan 10 items gemaakt zijn in de subtest.
- Er worden maximaal 20 items per subtest aangeboden bij non-verbale subtests. Bij verbale subtests zijn dit er maximaal 30.
- De subtest wordt afgebroken wanneer de standaardmeetfout van de schatting van de vaardigheid onder de 0,60 uitkomt.

In figuur 1 is een voorbeeld weergegeven van een adaptieve testprocedure (Universiteit van Amsterdam, 2010). Door telkens een volgend item aan te bieden met de meeste informatie, wordt de test aangepast aan het niveau van de leerling en ontstaat er een adaptieve test. De basis van de IRT-theorie en het adaptieve karakter van de ADIT wordt gebruikt voor het begrijpen van de ADIT en het analyseren van de gegevens. Meer informatie over de manier waarop de onderzoeksgegevens in dit onderzoek worden geanalyseerd is te vinden in hoofdstuk 3.



Figuur 1. De werking van Computer Adaptief Testen in het licht van de moderne testtheorie. Overgenomen uit @daptief toetsen, door de Universiteit van Amsterdam, 2010, p. 4, <http://www.uva.nl/binaries/content/documents/personalpages/k/l/s.klinkenberg/en/tab-two/tab-two/cpitem%5B11%5D/asset?1355372527110>

2.5. Waarom adaptief testen?

Een voordeel van adaptief testen is dat de items beter aansluiten bij het niveau van de leerling. Hierdoor wordt voorkomen dat de leerling items aangeboden krijgt met een subjectieve kans van $p < 0,5$ of $p > 0,5$. Items die niet passen bij het niveau zeggen immers weinig over zijn of haar capaciteiten en kunnen daarom weggelaten worden. Theoretisch gezien zal dit tot minder gokgedrag leiden, wat weer bijdraagt aan de betrouwbaarheid van de schatting. Indien de kwaliteit van de items (met name het discriminerend vermogen) voldoende is, kan de test ook met minder items afgenomen worden (van Zand, Pelt & Schrijver, 2015). Dit heeft als voordeel dat er sprake is van een kortere testduur. Schoonman (2011) spreekt hierbij echter van een productiviteitsparadox; een groot deel van de gewonnen tijd gaat bij adaptief toetsen weer verloren doordat de kandidaat met steeds moeilijkere items geconfronteerd wordt. Deze confrontatie van steeds moeilijkere vragen kan ook tot gevolg hebben dat adaptief toetsen, in tegenstelling tot wat er in hoofdstuk 1 beweerd is, wel degelijk tot frustratie bij leerlingen kan leiden. Bij adaptief toetsen verdwijnt namelijk de voldoening die ontstaan is wanneer de leerling zeker weet dat hij/zij een aantal items juist heeft beantwoord. Daarnaast zijn adaptieve toetsen bijna altijd digitaal en direct beoordeelbare toetsen. Vandaar dat het ook de nadelen heeft van een gesloten toets die voornamelijk bestaat uit meerkeuzevragen. Open vragen en productieve vaardigheden zijn onderdelen die nog niet met een adaptieve test gemeten kunnen worden. Daarnaast werken adaptieve toetsen vaak met 50% scores: de helft van de vragen moeten correct worden beantwoord en de helft incorrect. Onderzoek wijst echter aan dat het motiverender werkt wanneer het overgrote deel van de vragen correct beantwoord kan worden (Houtveen & Reezigt, 2000). Daarnaast is de omvang van de itembank een kritisch punt. Deze moet namelijk bestaan uit een groot aantal items. Weiss (1985) geeft aan dat itembanken die dienen voor adaptief testen ten minste uit 150 à 200 items moeten bestaan. Ook moeten deze items na een bepaalde tijd of na een bepaald aantal keren gebruikt te zijn uit de itembank gehaald worden om te voorkomen dat items bekend raken (Joosten-ten Brinke, Weber & Saveski, 2016).

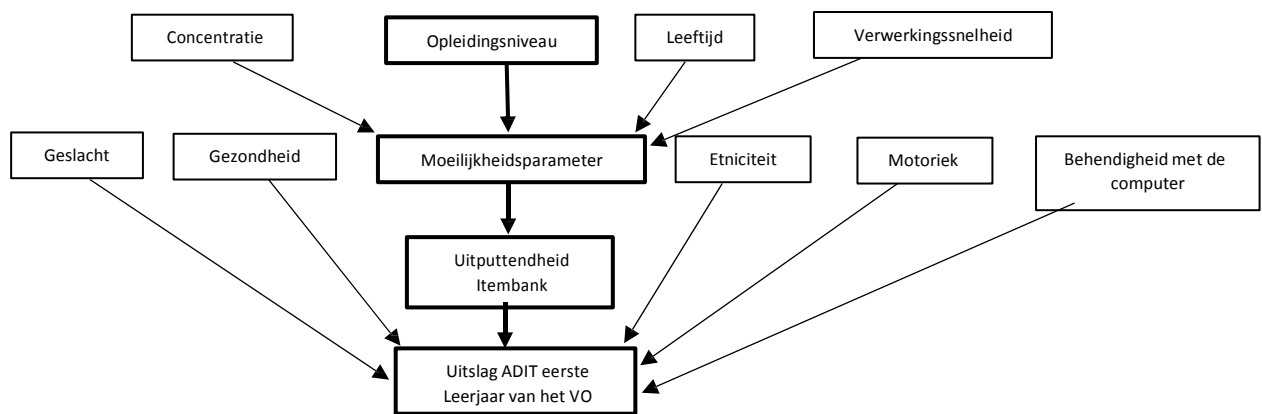
2.6. De ADIT

De ADIT wordt aangeboden aan leerlingen uit groep 8. Voor leerlingen uit groep 6 en 7 van de basisschool is de ADIT in ontwikkeling. De Cattell-Horn-Carroll-theorie (CHC-theorie) heeft als basis gediend voor de ontwikkeling van de ADIT. Deze theorie is ontstaan in 1990 door McGrew die de Gf-Gc Theory en Carroll's Three-Stratum Theory samen heeft gevoegd in een poging de verschillen tussen de twee theorieën te constateren (Alfonso, Flanagan & Radwan, 2005; Resing & Drenth, 2007). Bij de ontwikkeling van de ADIT is uitgegaan van een algemene intelligentiefactor (g). Deze wordt gevormd door twee onderliggende factoren: een verbale factor en een non-verbale factor. De verbale factor bestaat uit de subtesten: Woorden Passen (WP), Woordbetekenissen (WB) en Woordexclusie (EX). De non-verbale factor bestaat uit de subtesten: Kubussen (KB), Vlakken Leggen (VL), Reeksen (RK) en Hoofdbewerkingen (HB). Het onderscheid verbaal/non-verbaal vloeit voort uit het gehanteerde intelligentiemodel van Carroll (Vegt & Metselaar, 2013). De testuitslagen worden gegeven in de vorm van een Totale Intelligentie Score (TIS), een Verbale Intelligentie Score (VIS) en het Non-Verbale Intelligentie Score (NIS), met bijbehorende betrouwbaarheidsintervallen (Vegt & Metselaar, 2013).

2.7. Conceptueel model

Cognitieve vaardigheden van individuen ontwikkelen zich door tot in de adolescentie. Naarmate men ouder wordt zijn individuen dus, gemiddeld genomen, steeds beter in staat moeilijkere en complexere cognitieve problemen op te lossen. Hypothetisch zou gesteld kunnen worden dat de huidige ADIT-itembank over te weinig items met een hogere moeilijkheidsparameter beschikt die goed passend zijn bij het niveau van de nieuwe doelgroep (subjectieve kans $< 0,5$). In dit geval is er dus sprake van een discrepantie tussen het niveau van de aangeboden vragen en het niveau van de leerlingen in het eerste leerjaar van het voortgezet onderwijs. In dat verlengde zou hypothetisch gesteld kunnen worden dat de ADIT in zijn huidige staat niet geïmplementeerd kan worden in de brugklas. Onder het conceptueel model (figuur 2) is voor de volledigheid een overzicht weergegeven van de opgestelde hypothesen bij dit onderzoek.

Omdat meerdere factoren van invloed zijn op de uitslag van de ADIT, is de onderlinge samenhang van deze factoren weergegeven in een conceptueel model, zoals die in figuur 2 te vinden is. Zo is te zien dat de uitslag van de ADIT door de leerlingen vanuit het eerste leerjaar van het VO beïnvloed wordt door een groot aantal variabelen, waaronder de uitputtendheid van de itembank. Maar ook de gezondheid van de leerling en de behendigheid met de computer kunnen van invloed zijn op de uitslag. Met dikke lijnen en vlakken is aangegeven naar welke factoren en onderlinge causaliteit er in dit onderzoek gekeken wordt. Deze variabelen sluiten het beste aan bij de hoofdvraag van dit onderzoek.



Figuur 2. Conceptueel model

Voor dit onderzoek zijn verder de volgende hypothesen opgesteld:

- Hypothese 1: De ADIT bevat voldoende items met een hoge moeilijkheidsparameter om aan te sluiten op de verschillende, en met name hogere, niveaus wanneer $\leq 10\%$ van de leerlingen 50% van de 25% moeilijkste vragen aangeboden krijgen.
- Hypothese 2: Wanneer $\leq 2,5\%$ van de leerlingen een TIQ behaald van hoger dan 130 is er geen sprake van discrepantie tussen het niveau van de aangeboden vragen en het niveau van de leerlingen in het eerste leerjaar.
- Hypothese 3: Wanneer $\geq 75\%$ van de respondenten “mee eens” of “helemaal mee eens” op de stelling: “De testafname van de ADIT vond ik positief” aangeeft, wordt aangenomen dat de afname van de ADIT door de respondenten als positief wordt ervaren.
- Hypothese 4: De ADIT kan in zijn huidige staat geïmplementeerd worden in de brugklas, wanneer:
- $\leq 10\%$ van de leerlingen 50% van de 25% moeilijkste vragen aangeboden krijgen.
 - $\leq 2,5\%$ van de leerlingen een TIQ behaald hoger dan 130.
 - $\geq 75\%$ van de respondenten “mee eens” of “helemaal mee eens” op de stelling “De testafname van de ADIT vond ik positief” aangeeft.

3. Methoden van onderzoek

In dit hoofdstuk wordt de onderzoeksopzet en de hiervoor benodigde instrumenten beschreven. Vervolgens wordt er een beeld gegeven van de doelgroep, de onderzoeksprocedure en de analyses. Tot slot wordt er stil gestaan bij de betrouwbaarheid en validiteit van de onderzoeksopzet.

3.1. Onderzoeksopzet

Dit onderzoek naar de implementatie van de ADIT in het eerste leerjaar van het voortgezet onderwijs is zowel beschrijvend als evaluerend van aard (Van der Zee, 2005). Zo wordt er onderzocht of de huidige itembank van de ADIT geschikt is voor eerstejaars leerlingen (beschrijvend). Ook worden de leerlingen gevraagd naar hun mening over en ervaring met de ADIT (evaluerend). Kwantitatieve onderzoeksmethoden worden ingezet om gegevens te verzamelen. Zo wordt een enquête ingezet om de ervaring en de mening van de leerlingen over de ADIT te bevragen en wordt de ADIT als intelligentietest ingezet om informatie over de doelgroep te verzamelen. Dit betreft informatie met betrekking tot de in hoofdstuk 2 benoemde factoren: uitslag van de ADIT, leeftijd en geslacht van de leerlingen.

3.2. Doelgroep

De uitspraken van dit onderzoek zullen betrekking hebben op havo en vwo-leerlingen uit het eerste leerjaar van het voortgezet onderwijs. De doelgroep van dit onderzoek zijn echter havo en vwo-leerlingen uit het eerste en tweede leerjaar van het voortgezet onderwijs. Omdat dit onderzoek aan het begin van het schooljaar afgenomen wordt, zou de tijd die de leerlingen op het VO hebben doorgebracht te kort zijn om te spreken van een mogelijk intelligentieverschil met eind groep 8 leerlingen. Door aan het begin van het schooljaar tweedejaars leerlingen als doelgroep te kiezen is ook de opgedane kennis van leerjaar 1 meegenomen. Dit kan leiden tot een intelligentieverschil (crystallized intelligence). Omdat de intelligentietest geschikt moet zijn voor het gehele eerste leerjaar (van het begin tot het eind), wordt de ADIT in dit onderzoek afgenomen onder zowel leerlingen uit het eerste als leerlingen uit het tweede leerjaar. In hoofdstuk 4 komt de spreiding van leerlingen uit het eerste en tweede leerjaar in dit onderzoek aan bod. De gemiddelde leeftijd van de populatie (leerjaar 1 leerlingen VO) is 12 jaar. De verdeling man en vrouw is ongeveer gelijk (51,3% man en 48,7% vrouw). Het grootste deel van de populatie volgt onderwijs in West-Nederland (46.3%), gevolgd door Oost-Nederland (22.9%), Zuid-Nederland (20.3%) en Noord-Nederland (10.6%). Alle percentages met betrekking tot de populatie zijn terug te vinden in bijlage B2 (Centraal Bureau voor de Statistiek, 2016a, 2016b). In hoofdstuk 4 worden de gegevens m.b.t. de steekproef van dit onderzoek op dezelfde manier weergegeven.

Er wordt in dit onderzoek gebruik gemaakt van een getrapte steekproef (clustersteekproef en praktisch bruikbaar). Er wordt getracht een zo representatief mogelijke steekproef in dit onderzoek te gebruiken. Met een populatie van 21.593 eerstejaars VO-leerlingen is dit echter voor dit onderzoek niet mogelijk. Er zou een steekproef van duizend leerlingen nodig moeten zijn om de steekproef representatief voor de populatie te maken (Tellegen, 2004). Rekening houdend met een verwachte respons van 40% moeten er minimaal 32 scholen benaderd worden ($((500/40) * 100) : 40 = 32$). In totaal zijn er in dit onderzoek 158 scholen benaderd waarvan er 8 scholen meedoen met in totaal 406 proefpersonen ($n = 406$). Deze steekproef is groot genoeg om dit onderzoek te laten dienen als pilot of verkennend onderzoek naar de mogelijkheid van de ADIT in het eerste leerjaar van het VO. Er zijn zowel scholen benaderd die al bekend zijn met A-VISION evenals scholen die nog niet bekend zijn met het bedrijf.

3.3. Onderzoeksinstrumenten -ADIT

De ADIT is een digitale test en zal online op computers gemaakt moeten worden. De testafname verloopt via het testprogramma dat klaar staat op de server van A-VISION (Routewijs). Om de kans op problemen met de software en het online testplatform zo beperkt mogelijk te houden en om er zeker van te zijn dat de ICT-omgeving van de school geschikt is voor afname, doorlopen de scholen voorafgaand aan de afname een zogenoemd 'ben ik geschikt'-stappenplan. Dit stappenplan is als bijlage 5 toegevoegd. Om de onderzoeksvragen te kunnen beantwoorden, wordt er gebruik gemaakt van de huidige ADIT-itembank. Deze itembank is in het bezit van A-VISION. In paragraaf 3.5 is terug te vinden hoe de resultaten geanalyseerd worden.

3.4. Onderzoeksinstrumenten -Enquête

Aansluitend aan de intelligentietest wordt er bij de leerlingen een korte enquête afgenomen. De enquête zal eveneens digitaal via Routewijs worden aangeboden. De enquête zal bestaan uit zestien vragen en het invullen zal de leerling vijf minuten kosten. De afname van de intelligentietest duurt anderhalf uur. Om de bereidheid voor het invullen van de enquête te vergroten is er gekozen voor een korte enquête. Door middel van de enquête wordt beoogd een beeld te schetsen van de ervaring van de leerlingen met de ADIT. Omdat er geen dergelijke enquête over de ervaring met de ADIT bestaat, is dit tijdens dit onderzoek zelf ontwikkeld. Na overleg met de opdrachtgever is de uiteindelijke enquête tot stand gekomen (A. Egberink, A. Lubbers & C. Janssen, persoonlijke communicatie, 14 september, 2016). De enquête bestaat uit vijf schalen: algemene ervaring, ervaring met betrekking tot de lengte van de test, ervaring met betrekking tot de vraagstelling, ervaring met betrekking tot de vormgeving en ervaring met betrekking tot het niveau van de test. Deze schalen zijn in overleg met de opdrachtgever tot stand gekomen en trachten de gehele ervaring met de ADIT te ondervangen. De enquête is te vinden in bijlage B6. De respondenten worden gevraagd aan te geven in welke mate ze akkoord gaan met een uitspraak op een 4-puntsschaal, lopend van 1 (niet mee eens) tot 4 (helemaal mee eens). Hierbij geldt: hoe hoger de score, hoe positiever de ervaring met de ADIT. Met de afname van deze enquête wordt beoogd antwoord te geven op hypothese 3. Wanneer $\geq 75\%$ van de respondenten “mee eens” of “helemaal mee eens” aangeeft op de stelling: “De testafname van de ADIT vond ik positief”, wordt aangenomen dat de afname van de ADIT door de respondenten als positief wordt ervaren. De gegevens uit de enquête worden verzameld en ingevoerd in SPSS. Via een frequentieanalyse wordt getracht antwoord te geven op hypothese 3.

3.5. Onderzoeksprocedure

Het doel is om de ADIT bij een zo groot mogelijke doelgroep af te nemen. Om dit te bereiken wordt de doelgroep benaderd door via de mail of telefonisch contact op te nemen met zorgcoördinatoren van scholen. Dit gebeurt aan de hand van een standaard mail. Deze is als bijlage aan dit verslag toegevoegd (zie bijlage B3). Scholen kunnen de ADIT afnemen tussen november 2016 en eind januari 2017. Wanneer scholen toezeggen mee te doen aan het onderzoek zal een aanmeldformulier worden toegestuurd waarop ze verzocht worden leerlinggegevens in te vullen. Hierbij zal worden gevraagd naar de naam van de leerling, het geslacht, de geboortedatum, de klas en het niveau, of er sprake is van dyslexie en of de leerling óf een van de ouders in het buitenland is geboren. Deze gegevens worden verzameld voor gebruik tijdens en na het onderzoek. Zoals in hoofdstuk 2 is aangegeven, zullen de resultaten aan de hand van de klas en het niveau van de leerling worden geanalyseerd. Wanneer dit aanmeldformulier weer retour wordt gestuurd, worden er bij A-VISION per leerling inloggegevens aangemaakt (een gebruikersnaam en wachtwoord). Met de inloggegevens kunnen de leerlingen via routewijs bij de ADIT en aansluitende enquête. In bijlage B8 is een schematisch overzicht te vinden van de afnameprocedure bij A-VISION. Scholen mogen zelf kiezen of ze de test zelfstandig willen afnemen of dat ze willen dat de testafname begeleid wordt door een medewerker vanuit A-VISION. Ondersteuning van docenten of een testbegeleider is echter van belang bij vragen van leerlingen. Voorafgaand aan elk testonderdeel verschijnt de testinstructie op het beeldscherm. Ook moeten de leerlingen voor de start van elk testonderdeel eerst oefenopdrachten maken om te zien of ze de uitleg daadwerkelijk begrepen hebben. Wanneer de bedoeling van de opdracht onduidelijk is kan de docent of de, indien aanwezige, testbegeleider dit ophelderen. Om de standaardisatie van het onderzoek te vergroten krijgen de docenten allemaal dezelfde testinstructie toegestuurd. Wanneer er een testbegeleider aanwezig is, geeft deze ook dezelfde testinstructie. In de bijlage (B4) is de testinstructie toegevoegd. Na afloop van de test worden de gegevens verzameld en verwerkt via een Excel-bestand.

3.6. Analyses

Intelligentiescores zijn in een samenleving normaal verdeeld. Gemiddeld hebben de meeste mensen een intelligentiescore rond de 100, met een standaarddeviatie (SD) van 15. Gemiddeld behaald 68% van de leerlingen een intelligentiescore van beneden- tot bovengemiddeld niveau (-1 SD. ; +1 SD.) De meeste havo-leerlingen behalen een IQ-score tussen de 100 en 125. De meeste vwo-leerlingen behalen een IQ-score tussen de 110 en 135. De huidige ADIT zou dezelfde percentages moeten halen wil de itembank geschikt zijn voor implementatie in het eerste leerjaar. Items met een hogere moeilijkheidsparameter correleren met een hogere IQ-score (Embretson & Reise, 2000; Lai, et al., 2011). Vanuit deze beredenering wordt er tijdens dit onderzoek geanalyseerd met behulp van de moeilijkheidsparameters behorend bij de items. Om te onderzoeken of de huidige ADIT-itembank over voldoende items met een hoge moeilijkheidsparameter ($B > 1$) beschikt, zijn de items o.b.v. hun moeilijkheidsparameter grofweg in vier procentuele groepen ingedeeld. Vervolgens is er gekeken naar het percentage van de leerlingen waarvan, op basis van de normaalverdeling, verwacht mag worden dat ze de items met $B > 1$ aangeboden hebben gekregen. Op basis hiervan wordt vervolgens vastgesteld hoeveel procent van de 25% items met de hoogste moeilijkheidsparameters aangeboden mogen worden aan dat percentage leerlingen. Deze percentages en de beredenering achter deze analyse is bepaald in overeenstemming met de opdrachtgever.

Op basis hiervan is hypothese 1 als volgt opgesteld: De ADIT bevat voldoende items met een hoge moeilijkheidsparameter om aan te sluiten op de verschillende, met name hogere, niveaus wanneer $\leq 10\%$ van de leerlingen 50% van de 25% items met de hoogste moeilijkheidsparameter aangeboden krijgen. In Excel wordt per leerling weergegeven of een item is aangeboden (1)

of niet is aangeboden (0). Met behulp van Excel wordt er per leerling uitgerekend hoeveel van deze 25% items er aangeboden zijn. De moeilijkheidsparameters verschillen per subtest. Ook verschilt het aantal items dat behoort tot de 25% per subtest. Gemiddeld zijn dit 10 items, waarbij de moeilijkheidsparameter loopt in een range van 0,5269 tot 2,82727. Op deze manier kan er uitgerekend worden hoeveel procent van de leerlingen 50% van de 25% items met de hoogste moeilijkheidsparameter aangeboden hebben gekregen.

Naast het feit dat de items niet te makkelijk of te moeilijk moeten zijn, dient dit onderzoek ook als opzet voor een vervolg- /normeringsonderzoek. Om deze reden wordt er bij de analyses ook gekeken naar de mogelijkheid tot normering van de ADIT voor het eerste leerjaar van het voortgezet onderwijs. Om een betrouwbare normering te krijgen moet er sprake zijn van een afstemming tussen het niveau van de aangeboden vragen en het niveau van de proefpersoon. Omdat dit onderzoek havo en vwo-leerlingen als doelgroep heeft, zou men verwachten dat de IQ-scores van deze groep vallen tussen de 95 en de 125 (Resing & Drenth, 2007). Grote uitval naar boven of beneden ten opzichte van de verwachte TIQ-score (95-125) zouden kunnen worden geïnterpreteerd als gevolg van een discrepantie tussen het niveau van de aangeboden vragen en het niveau van de respondenten (van Minden, 2010; Resing & Drenth, 2007). Op basis van deze gegevens is hypothese 2 opgesteld: Wanneer $\leq 2,5\%$ van de leerlingen een TIQ behaald van hoger dan 130 is er geen sprake van discrepantie tussen het niveau van de aangeboden vragen en het niveau van eerste jaars VO leerlingen. Hierbij moet worden opgemerkt dat er gerekend wordt met testcores die genormeerd zijn voor groep 8 leerlingen. De ADIT zal eerst moeten worden genormeerd voor de VO doelgroep om uiteindelijk ingezet te kunnen worden voor leerjaar 1 leerlingen. Toch is er gekozen om te werken met deze testcores omdat er in dit onderzoek gewerkt wordt met de huidige ADIT-itembank en dus ook de huidige moeilijkheidsparameters. Op basis van de literatuur kan hypothetisch gesteld worden dat er te weinig items met een moeilijkheidsparameter van >2 in de huidige itembank aanwezig zijn. Vanuit de literatuur weten we dat items met een hogere moeilijkheidsparameter correleren met een hogere IQ-score (Embretson & Reise, 2000; Lai, et al., 2011). Op basis hiervan kan er hypothetisch worden gesteld dat uit het onderzoek meer leerlingen komen met een hogere IQ-score (hier is gekozen voor $IQ > 130$). Vanuit deze beredenering, puur gericht om te kijken naar het bereik van de moeilijkheidsparameters van de items, is er toch voor gekozen om de voor groep 8 genormeerde IQ-scores in de analyses mee te nemen.

Met dit onderzoek wordt beoogd antwoord te geven op de hoofdvraag zoals beschreven in hoofdstuk 2. Passend bij deze hoofdvraag is de volgende hypothese (4) opgesteld: de ADIT kan in zijn huidige staat geïmplementeerd worden in de brugklas. Fase 5 van het proces van testontwerp en testontwikkeling is de testrevisie. Revisie betekent het controleren en eventueel corrigeren. Revisie van de test hoeft niet bij elke testontwikkeling gedaan te worden. Revisie is alleen belangrijk wanneer er bijvoorbeeld nieuwe wetenschappelijke inzichten zijn, gedateerd testmateriaal, verandering van populatie of curriculum, etcetera. Omdat er bij dit onderzoek gekeken wordt naar een andere populatie, moet er dus een testrevisie plaatsvinden (de Zeeuw, 2000). Er moet worden gekeken of de ADIT met een veranderende doelgroep nog altijd inzetbaar is. De hypothese (4) zal worden aangenomen wanneer hij voldoet aan de volgende punten:

- Wanneer 10% van de leerlingen 50% van de 25% moeilijkste vragen aangeboden krijgen.
- Wanneer $\leq 2,5\%$ van de leerlingen een TIQ behaald hoger dan 130.
- Wanneer $\geq 75\%$ van de respondenten “mee eens” of “helemaal mee eens” op de stelling “De testafname van de ADIT vond ik positief” aangeeft.

3.7. Betrouwbaarheid en validiteit onderzoeksdesign- en instrumenten

Dit onderzoek heeft als doel een betrouwbare en valide uitspraak te kunnen doen met betrekking op de gehele onderzoekspopulatie. Om dit te realiseren is er gebruik gemaakt van een getrapte steekproef. Om de representativiteit en daarmee de generaliseerbaarheid van de uitspraken te vergroten zijn er in totaal 158 scholen benaderd, als volgt verspreid over Nederland: Zuid (18), Oost (41), Noord (33) en West (66). Dit is gebeurd in overeenkomst met het totale aantal leerlingen per regio (zie tabel 16, bijlage B2). In paragraaf 5.5 wordt nader ingegaan op de generaliseerbaarheid en representativiteit van dit onderzoek. Om de betrouwbaarheid zo goed mogelijk te waarborgen is het onderzoek zo transparant mogelijk uitgevoerd en zijn er maatregelen genomen om de standaardisatie te verhogen. Zo zijn er voorafgaand aan de afname voor de enquête standaardvragen ontwikkeld die op dezelfde manier aan de respondenten zijn aangeboden. Bij het opstellen van de enquêtevragen is er zoveel mogelijk gebruik gemaakt van een Likertschaal. Om te waarborgen dat de vragen duidelijk, helder en eenduidig beschreven zijn, is de enquête daarna onderworpen aan een proefafname. Tijdens dit onderzoek wordt er gebruik gemaakt van een COTAN voldoende beoordeelde test: de ADIT. Om ervoor te zorgen dat de ICT-omgeving van de

scholen gelijk is, moet de school voor afname een 'ben ik geschikt'-test doen. Toch sluit dit eventuele technische problemen op de dag zelf niet uit. Hier moet rekening mee worden gehouden bij het interpreteren van de resultaten en betrouwbaarheid van dit onderzoek. De testinstructies worden op het beeldscherm van de leerling weergegeven en zijn voor eenieder gelijk. De betrouwbaarheid van het onderzoek wordt verlaagd door het feit dat de afname van de ADIT niet gestandaardiseerd is en de testleider verschilt per school. Dit kan tot gevolg hebben dat een klas bijvoorbeeld meer last heeft van ruis tijdens de afname of dat een testleider meer bekwaam is in het beantwoorden van vragen van leerlingen. Deze factoren kunnen van invloed zijn op de betrouwbaarheid van het onderzoek. Hier wordt in hoofdstuk 5 uitgebreid bij stilgestaan.

4. Onderzoeksresultaten

In dit hoofdstuk zal het verloop van en de respons op het onderzoek worden besproken. Daarnaast vindt u hier de onderzoeksresultaten per deelvraag weergegeven.

4.1. Respons en uitvoering- ADIT

Scholen konden de ADIT afnemen tussen november en januari. In de maand november is er met de meeste scholen telefonisch contact geweest en werden de eerste afspraken voor het onderzoek gemaakt. In de maand januari werd de ADIT bij de meeste leerlingen afgenomen. Er was mogelijkheid tot uitloop van de testafname tot midden februari. In totaal is de ADIT bij 406 (n=406) leerlingen afgenomen. Leerlingen die de ADIT niet helemaal hebben afgerond omdat ze ziek of afwezig waren, zijn in dit onderzoek niet meegenomen. Tevens zijn leerlingen met een testscore onder de 80 ook niet meegenomen, omdat deze leerlingen de ADIT niet serieus hebben ingevuld. Van de proefpersonen is 50% man en 50% vrouw. De gemiddelde leeftijd van de van de proefpersonen bedraagt 12,9 jaar en varieert tussen 11 en 15 jaar (sd = 0,8). Het merendeel van de proefpersonen is afkomstig uit de regio Oost (69,7%), gevolgd door West (22,4%) en Noord (7,9%). In tabel 1 zijn verdere gegevens van de steekproef uitgewerkt.

Tabel 1 Sekse en opleidingsniveau van de respondenten

	Man	Vrouw	Totaal	% van de totale respons ADIT (%n)
Leerjaar 1				
Havo	35	53	88	21,7%
Havo/vwo	63	55	118	29,1%
Vwo	30	48	78	19,2%
Totaal	128	156	284	70%
Leerjaar 2				
Havo	49	32	81	20%
Vwo	26	15	41	10%
Totaal	75	47	122	30%
Totale respons (n)	203	203	406	100%

4.2. Respons en uitvoering- Enquête

Ook is er bij de respondenten een enquête afgenomen. Door tijdsgebrek of onduidelijkheden is de enquête niet door alle respondenten ingevuld. In totaal is de enquête door 358 respondenten ingevuld (n=358). Niet afgeronde enquêtes zijn in dit onderzoek niet meegenomen. Van de respondenten is 48,9% man en 51,1% vrouw. De gemiddelde leeftijd van de respondenten bedraagt 12,9 jaar en varieert tussen 11 en 15 jaar (sd = 0,8). Het merendeel van de respondenten is afkomstig uit de regio Oost (78,5%), gevolgd door West (12%) en Noord (9,5%). In tabel 2 zijn verdere gegevens van de respondenten uitgewerkt.

Tabel 2 Sekse en opleidingsniveau van de respondenten

	Man	Vrouw	Totaal	% van de totale respons ADIT (%n)
Leerjaar 1				
Havo	36	53	89	24,9%
Havo/vwo	52	47	99	27,6%
Vwo	31	48	79	22,1%
Totaal	119	148	267	74,6%
Leerjaar 2				
Havo	44	27	71	19,8%
Vwo	12	8	20	5,6%
Totaal	56	35	91	25,4%
Totale respons (n)	175	183	358	100%

4.3. Resultaten deelvraag 1

In hoeverre bevat de huidige ADIT-itembank voldoende items om een door A-VISION vooraf bepaalde, nauwkeurige TIQ-meting te geven voor leerlingen uit het eerste leerjaar havo/vwo van het VO?

Om deelvraag 1 te kunnen beantwoorden is er per subtest een analyse gemaakt in Excel. De resultaten van die analyse worden per subtest besproken. Per subtest is 25% van de items met de hoogste moeilijkheidsgraad geselecteerd. Zoals in hypothese 1 wordt beschreven mag maar 10% van de leerlingen 50% van deze moeilijkste items aangeboden krijgen, om te kunnen stellen dat de huidige ADIT-itembank over voldoende items beschikt. In Excel is er daarom een analyse gemaakt waarin dit percentage wordt berekend. Er wordt bij de analyse een onderscheid gemaakt tussen proefpersonen uit leerjaar 1 en 2, evenals het niveau van de proefpersonen. Gevonden verschillen zijn geanalyseerd door middel van een t-toets voor onafhankelijke steekproeven met een alfa van 5%.

Op het onderdeel Woord Exclusie wordt > 50% van de 25% moeilijkste items aangeboden aan in totaal 87 leerlingen (21,4%). Er is een verschil gevonden tussen havo en vwo-leerlingen uit zowel leerjaar 1 als leerjaar 2. Daarnaast is er ook een verschil gevonden tussen leerjaar 1 en leerjaar 2 leerlingen, ongeacht het niveau. De resultaten zijn terug te vinden in onderstaande tabel. Uit de t-toets bleek dat het verschil tussen de scores van leerjaar 1 leerlingen en leerjaar 2 leerlingen statistisch significant is ($p = 0,01$, $n = 406$). Een significant verschil is ook gevonden tussen havo 1 en vwo 1 leerlingen ($p < 0,01$, $n = 284$) en havo 2 en vwo 2 leerlingen ($p < 0,01$, $n = 122$).

Tabel 3 Aantal leerlingen op het onderdeel Woord Exclusie die >50% aangeboden heeft gekregen; op leerjaar en niveau

Woord Exclusie	Havo	% van totaal havo*	Havo/vwo	% van totaal havo/vwo*	Vwo	% van totaal vwo*	Totaal	% van het totaal*
Leerjaar 1	8	9%	15	12,7%	31	39,7%	54	19%
Leerjaar 2	14	17,3%			19	46,3%	33	27%
Totaal	22	13%	15	12,7%	50	42%	87	21,4%

* Het totaal van de bijbehorende rij

Op het onderdeel Hoofdbewerkingen wordt >50% van de 25% moeilijkste items aangeboden aan in totaal 52 leerlingen (12,8%). In onderstaande tabel zijn gegevens per leerjaar en niveau uitgewerkt. Hierin is te zien dat er ook op dit onderdeel verschillen gevonden zijn tussen leerlingen uit de verschillende niveaus en leerjaren. Uit de t-toets bleek dat het verschil tussen de score van leerjaar 1 leerlingen en leerjaar 2 leerlingen statistisch significant is ($p < 0,01$, $n = 406$). Ook tussen havo en vwo-niveau is er in beide leerjaren een statistisch significant verschil gevonden. In leerjaar 1 is dit verschil statistisch significant met $p < 0,01$ ($n = 284$). In leerjaar 2 is dit verschil statistisch significant met $p < 0,01$ ($n = 122$).

Tabel 4 Aantal leerlingen op het onderdeel Hoofdbewerkingen die >50% aangeboden heeft gekregen; op leerjaar en niveau

Hoofdbewerkingen	Havo	% van totaal havo*	Havo/vwo	% van totaal havo/vwo*	Vwo	% van totaal vwo*	Totaal	% van het totaal*
Leerjaar 1	8	9%	9	7,6%	14	17,9%	31	10,9%
Leerjaar 2	9	11%			12	29,3%	21	17,2%
Totaal	17	10%	9	7,6%	26	21,8%	52	12,8%

* Het totaal van de bijbehorende rij

Op het onderdeel Vlakken wordt >50% van de 25% moeilijkste items aangeboden aan in totaal 41 leerlingen (10%). Uit tabel 5 is te herleiden dat onder leerjaar 1 en leerjaar 2 leerlingen, evenals onder havo en vwo-leerlingen in beide leerjaren verschillen zijn gevonden in de frequentie van de aangeboden items. Het gevonden verschil bleek statistisch significant te zijn ($p < 0,01$, $n = 406$). In beide leerjaren is er tussen havo en vwo ook een statistisch significant verschil gevonden, waar deze in het eerste leerjaar een p-waarde heeft van $p < 0,01$ ($n = 284$). In het 2^e leerjaar is er een p-waarde van $p < 0,01$ gevonden ($n = 122$).

Tabel 5 Aantal leerlingen op het onderdeel Vlakken die >50% aangeboden heeft gekregen; op leerjaar en niveau

Vlakken	Havo	% van totaal havo*	Havo/vwo	% van totaal havo/vwo*	Vwo	% van totaal vwo*	Totaal	% van het totaal*
Leerjaar 1	3	3,4%	12	10%	11	14,1%	26	9,2%
Leerjaar 2	7	8,6%			8	19,5%	15	12,3%
Totaal	10	5,9%	12	10%	19	16%	41	10%

* Het totaal van de bijbehorende rij

Meer dan de helft van de moeilijkste items van het onderdeel Woordenpassen is aan 14,3% van de respondenten voorgelegd. In onderstaande tabel zijn verdere gegevens van de steekproef uitgewerkt. Ook hier is er een verschil tussen het leerjaar en het niveau van de leerlingen gevonden. Uit de t-toets bleek dat het verschil tussen de scores van leerjaar 1 leerlingen en leerjaar 2 leerlingen statistisch significant is ($p < 0,01$, $n = 406$). Een statistisch significant verschil is ook gevonden tussen havo 1 en vwo 1 leerlingen ($p < 0,01$, $n = 284$) en havo 2 en vwo 2 leerlingen ($p < 0,01$, $n = 122$).

Tabel 6 Aantal leerlingen op het onderdeel Woorden Passen die >50% aangeboden heeft gekregen; op leerjaar en niveau

Woorden passen	Havo	% van totaal havo*	Havo/vwo	% van totaal havo/vwo*	Vwo	% van totaal vwo*	Totaal	% van het totaal*
Leerjaar 1	6	6,8%	11	9,3%	19	24,4%	36	12,7%
Leerjaar 2	9	11,1%			13	31,7%	22	18%
Totaal	15	8,9%	11	9,3%	32	26,9%	58	14,3%

* Het totaal van de bijbehorende rij

Bij het onderdeel Woord Betekenis wordt >50% van de 25% moeilijkste items aangeboden aan in totaal 28 leerlingen (6,9%). Overige gegevens zijn in onderstaande tabel (7) uitgewerkt. Ook op dit onderdeel zijn bovengenoemde verschillen gevonden. Uit de t-toets bleek dat het verschil tussen de scores van leerjaar 1 en leerjaar 2 leerlingen statistisch significant is ($p < 0,01$, $n = 406$). Ook tussen havo en vwo-niveau is er op beide leerjaren een statistisch significant verschil gevonden. In leerjaar 1 is dit verschil statistisch significant met $p < 0,01$ ($n = 184$). In leerjaar 2 is dit verschil statistisch significant met $p < 0,01$ ($n = 122$).

Tabel 7 Aantal leerlingen op het onderdeel Woord Betekenis die >50% aangeboden heeft gekregen; op leerjaar en niveau

Woord betekenis	Havo	% van totaal havo*	Havo/vwo	% van totaal havo/vwo*	Vwo	% van totaal vwo*	Totaal	% van het totaal*
Leerjaar 1	1	1,1%	3	2,5%	10	12,8%	14	4,9%
Leerjaar 2	7	8,6%			7	17,1%	14	11,5%
Totaal	8	4,7%	3	2,5%	17	14,3%	28	6,9%

* Het totaal van de bijbehorende rij

Op het onderdeel Kubussen wordt >50% van de 25% moeilijkste items aangeboden aan in totaal 44 leerlingen (10,8%). Zoals in tabel 8 terug te lezen is, is er een verschil gevonden onder leerlingen uit leerjaar 1 en leerjaar 2, evenals tussen leerlingen van verschillende niveaus. Het gevonden verschil tussen scores van eerste- en tweedejaars leerlingen bleek statistisch significant te zijn ($p < 0,01$, $n = 406$). In beide leerjaren is er tussen havo en vwo ook een statistisch significant verschil gevonden. Waar deze in het eerste leerjaar een p-waarde heeft van $p = 0,03$ ($n = 284$), is in het 2^e leerjaar een p-waarde van $p < 0,01$ gevonden ($n = 122$).

Tabel 8 Aantal leerlingen op het onderdeel Kubussen die >50% aangeboden heeft gekregen; op leerjaar en niveau

Kubussen	Havo	% van totaal havo*	Havo/vwo	% van totaal havo/vwo*	Vwo	% van totaal vwo*	Totaal	% van het totaal*
Leerjaar 1	3	3,4%	11	9,3%	11	14,1%	25	8,8%
Leerjaar 2	8	9,9%			11	26,8%	19	15,6%
Totaal	11	6,5%	11	9,3%	22	18,5%	44	10,8%

* Het totaal van de bijbehorende rij

Meer dan de helft van de moeilijkste items van het onderdeel Reeksen is aan 35% van de respondenten voorgelegd. De gegevens over dit onderdeel zijn terug te vinden in tabel 9. Ook op dit onderdeel is een statistisch significant verschil gevonden tussen de scores van leerjaar 1 leerlingen en leerjaar 2 leerlingen ($p < 0,01$, $n=406$). Evenals tussen de scores van havo en vwo-leerlingen in beide leerjaren. In leerjaar 1 is hierbij een p-waarde van $p < 0,01$ ($n= 284$) gevonden. In leerjaar 2 is er een statistisch significant verschil met een p-waarde van $p < 0,01$ ($n=122$) gevonden.

Tabel 9 Aantal leerlingen op het onderdeel Kubussen die >50% aangeboden heeft gekregen; op leerjaar en niveau

Reeksen	Havo	% van totaal havo*	Havo/vwo	% van totaal havo/vwo*	Vwo	% van totaal vwo*	Totaal	% van het totaal*
Leerjaar 1	18	20,5%	29	24,6%	36	46,2%	83	29,2%
Leerjaar 2	31	38,3%			28	68,3%	59	48,4%
Totaal	49	29%	29	24,6%	64	53,8%	142	35%

* Het totaal van de bijbehorende rij

Op basis van bovenstaande resultaten wordt hypothese 1 verworpen (De ADIT bevat voldoende items met een hoge moeilijkheidsparameter om aan te sluiten op de verschillende, met name hogere, niveaus wanneer $\leq 10\%$ van de leerlingen $\leq 50\%$ van de moeilijkste vragen aangeboden krijgen).

4.4. Resultaten deelvraag 2

In hoeverre bevat de huidige ADIT-itembank voldoende items met een hoge moeilijkheidsgraad ($B \geq 0$) om het intelligentieniveau van leerlingen uit het eerste leerjaar van het VO nauwkeurig in te schatten?

Cognitieve capaciteiten ontwikkelen zich tot in de adolescentie door (Resing & Drenth, 2007). De ADIT en zijn huidige itembank is voor groep 8 ontwikkeld. Dus zou hypothetisch gesteld kunnen worden dat er te weinig items met een hogere moeilijkheidsparameter van >1 in de huidige itembank aanwezig zijn om het niveau van leerlingen uit het eerste leerjaar van het VO nauwkeurig in te kunnen schatten. Wetende dat items met een hogere moeilijkheidsparameter correleren met een hogere IQ-score (Reise, 2000; Lai, et al., 2011), wordt er voor het beantwoorden van deelvraag 2 een analyse gemaakt over de, in dit onderzoek verzamelde, IQ-scores. Analyses zijn gemaakt over leerjaar 1 leerlingen, evenals leerjaar 2 en de totale steekproef. Om de scores inzichtelijker te maken zijn de IQ-scores, op basis van de normaalverdeling die in de samenleving voorkomt, tot groepen samengevoegd; (1). IQ-score ≤ 70 , (2). IQ-score tussen 71-85, (3). IQ-score tussen 86-115, (4). IQ-score tussen 116-129 en (5). IQ-score ≥ 130 . In tabel 10 zijn de resultaten van deze deelvraag overzichtelijk weergegeven.

De gemiddelde IQ-score van de gehele steekproef is 117, en varieert tussen 82 en 145 ($SD= 11$). In leerjaar 1 is er een gemiddelde IQ-score van 115 gevonden. De IQ-score varieert hierbinnen tussen 82 en 145 ($SD= 10,5$). In leerjaar 2 is er een gemiddelde IQ-score van 122 gevonden. Onder respondenten uit deze groep varieert de IQ-score tussen 95 en 145 ($SD= 11,2$). Uit de T-toets voor onafhankelijke steekproeven bleek dat met een alfa van 5% het verschil in IQ-scores tussen leerjaar 1 en leerjaar 2 statistisch significant is ($p < 0,01$, $n= 406$). De normaalverdelingen behorende bij bovenstaande gegevens zijn terug te vinden in bijlage 9. Bij het analyseren van de gegevens moet rekening worden gehouden met het feit dat voor de beantwoording van deelvraag 2 gewerkt is met IQ-scores die voor groep 8 leerlingen zijn genormeerd. Dit betekent dat, wanneer de scores voor de nieuwe doelgroep genormeerd zijn, er andere IQ-scores uit kunnen komen. De resultaten moeten dus met voorzichtigheid worden geïnterpreteerd. Daarnaast moet rekening worden gehouden met het feit dat in dit onderzoek alleen havo en vwo-leerlingen mee zijn genomen. Dit betekent dat er in dit onderzoek alleen gekeken wordt naar een gedeelte binnen de, voor de populatie bekende, normaalverdeling. Dit betekent ook dat de gevonden normaalverdeling

Tabel 10 Verdeling IQ-scores

		IQ-range	n	Percentage
Leerjaar 1		≤ 70	0	0%
		71-85	3	1,1%
		86-115	153	53,9%
		116-129	99	34,9%
	≥ 130	29	10,1%	
Leerjaar 2		≤ 70	0	0%
		71-85	0	0%
		86-115	41	33,6%
		116-129	51	41,8%
		≥ 130	30	24,6%
Totale steekproef		≤ 70	0	0%
		71-85	3	0,7%
		86-115	194	47,8%
		116-129	150	37%
		≥ 130	59	14,5%

binnen dit onderzoek afwijkt van de bekende normaalverdeling voor de populatie. Na dit gezegd te hebben, wordt voor het verwerpen of aannemen van de hypothese gekeken naar het percentage leerlingen die een IQ-score op de huidige ADIT laat zien van > 130. Op basis van de resultaten zoals gepresenteerd in tabel 10, wordt hypothese 2 verworpen (Wanneer $\leq 2,5\%$ van de leerlingen een TIQ behaald van hoger dan 130 is er geen sprake van discrepantie tussen het niveau van de aangeboden vragen en het niveau van de leerlingen in het eerste leerjaar).

4.5. Resultaten deelvraag 3

Ook is naar de ervaring van de proefpersonen met de ADIT gevraagd. Om te waarborgen dat de items van de enquête duidelijk, helder en eenduidig beschreven zijn, is de enquête voor afname onderworpen aan een proefafname. Na de proefafname is de homogeniteit van de vragenlijst met behulp van SPSS onderzocht. Om de betrouwbaarheid van de enquête te waarborgen is er gekeken naar de onderlinge samenhang van de items op iedere subschaal. De subschalen laten allemaal een Cronbach's alfa van >0,5 zien. Op de schaal tijdsduur wordt er een alfa gevonden van 0,631. Op de schaal vraagstelling is de alfa 0,779 en op de schaal vormgeving 0,575. Dit laat zien dat er voldoende onderlinge samenhang bestaat tussen de items om iets te kunnen zeggen over de ervaring van de leerlingen met de ADIT. Zoals in paragraaf 4.1 te lezen was, hebben sommige proefpersonen de enquête niet ingevuld. Omdat de steekproefgegevens verder wel met elkaar overeenkomen, kan ervan uit worden gegaan dat de steekproef van de enquête (n= 358) representatief is voor de gehele steekproef van dit onderzoek (n=406). Voor het beantwoorden van deelvraag 3 is het belangrijk om de frequentie van antwoorden op de vraag: "De testafname van de ADIT vond ik positief" te weten.

Deze vraag beantwoorde 58,7% van de respondenten met 'mee eens' (n= 210) en 6,7% met 'helemaal mee eens' (n= 24). Zo'n 65,4% van de respondenten was het dus eens met deze stelling en hebben de afname van de ADIT als positief ervaren. Op de individuele stellingen geven de respondenten gemiddeld aan dat ze het ermee eens zijn. Dit is met uitzondering van twee stellingen die betrekking hebben op de tijdsduur. 55,6% van de respondenten gaf aan dat ze onvoldoende tijd hadden om een vraag binnen de tijd in te vullen. Bij alle items is een spreiding gevonden tussen de 1 en 4. Overige gegevens met betrekking tot de items zijn te vinden in tabel 17, bijlage 9. Hieronder in tabel 11 zijn gegevens van drie items weergegeven. Deze items zijn hier weergegeven omdat ze van belang zijn bij het beantwoorden van deelvraag 3.

Tabel 11 Verdeling uitkomsten enquête						
	Frequentie	Percentage	Gemiddeld	STD	Modus	Range
De testafname vond ik positief			2,66	0,7	3	1-4
Niet mee eens	23	6,4%				
Een beetje mee eens	101	28,2%				
Mee eens	210	58,7%				
Helemaal mee eens	24	6,7%				
Hoe vond je de test?			2,04	0,415	2	1-3
Makkelijk	24	6,7%				
Voldoende op mijn niveau	296	82,7%				
Moeilijk	38	10,6				
Heeft de test je uitgedaagd?			1,27	0,638	1	1-3
Ja, voldoende	297	83%				
Te veel	24	6,7				
Te weinig	37	10,3				

Bij het analyseren van de gegevens is er ook gekeken naar het verschil, tussen havo en vwo-scholieren, in antwoorden op de drie bovengenoemde items. Uit de enkelvoudige variantieanalyse ANOVA bleek dat met een alfa van 5% het verschil in antwoorden tussen havo en vwo-scholieren, op deze drie items, niet significant te zijn ($p= 1,0$, $n= 358$). Omdat < 75% van de leerlingen 'mee eens' of 'helemaal mee eens' op de stelling "De testafname van de ADIT vond ik positief" antwoordde, wordt hypothese 3 verworpen (Wanneer $\geq 75\%$ van de respondenten 'mee eens' of 'helemaal mee eens' op de stelling "De testafname van de ADIT vond ik positief" aangeven, wordt aangenomen dat de afname van de ADIT door de respondenten als positief worden ervaren).

5. Conclusie, discussie, aanbevelingen

Vanuit de resultaten die in het vorige hoofdstuk gegeven zijn, wordt er antwoord gegeven op de onderzoeksvraag. Daarnaast wordt de betrouwbaarheid, validiteit en de bruikbaarheid van het onderzoek bediscussieerd en worden er aanbevelingen gedaan voor vervolgonderzoek.

5.1. In hoeverre bevat de huidige ADIT-itembank voldoende items om een, vooraf door A-VISION bepaalde, nauwkeurige TIQ-meting te geven voor leerlingen uit het eerste leerjaar havo/vwo van het VO?

Er is in dit onderzoek per subtest gekeken naar de moeilijkheidsparameter van de aangeboden vragen. Hierbij zijn er verschillen gevonden tussen subtesten. De subtesten vlakken en woordbetekenissen bleken wel geschikt te zijn om in te zetten voor leerlingen in het eerste leerjaar van het voortgezet onderwijs. De andere subtesten bleken hier niet geschikt voor te zijn. Ook zijn er significante verschillen gevonden tussen havo en vwo-leerlingen, evenals tussen leerjaar 1 en leerjaar 2 leerlingen. Uit het theoretisch kader is gebleken dat de kennis en de inzichten van leerlingen zich blijven ontwikkelen (Resing & Drenth, 2007). Deze theorie kan de gevonden verschillen tussen leerjaar 1 en leerjaar 2 leerlingen verklaren. Leerlingen uit leerjaar 2 hebben immers langer onderwijs gehad en hebben meer kennis tot hun beschikking dan leerlingen uit leerjaar 1. Het kan dan mogelijk ook dit verschil in onderwijstijd zijn die tot de gevonden verschillen hebben geleid. De gevonden verschillen tussen havo en vwo-leerlingen uit beide leerjaren zijn echter opvallender. Uit rapporten over de VOCL-bovenbouwstudies (Kuyper & Guldemon, 1997; Kuyper et al., 1999) komen enkele onderzoeksuitkomsten die dit verschil zouden kunnen verklaren. Zo blijkt uit bovengenoemd onderzoek dat de grootste gevonden verschillen tussen havo en vwo-leerlingen cognitieve verschillen betreffen. Met name het verschil in 'leervermogen' is een belangrijke factor. Havoleerlingen schijnen meer moeite te hebben met de zogenoemde 'transfer van kennis'; het toepassen van opgedane kennis in nieuwe theoretische situaties. Daarnaast is de mate van homogeniteit in havo/vwo-klassen een belangrijk gevonden verschil. In havoklassen is de spreiding in vaardigheden groter dan in vwo-klassen. Mogelijk is dit een verklaring voor de gevonden verschillen in dit onderzoek. De onderzochte vwo-leerlingen waren, vergeleken met de havoleerlingen, minder vertegenwoordigd. Een kleine groep met een kleine spreiding en een hoger gemiddelde resulteert in een grotere kans om meer vragen aangeboden te krijgen. Een andere mogelijke verklaring ligt op het gebied van de motivatie en concentratie. Over het algemeen zijn havisten meer gericht op een uitkomst, terwijl vwo'ers gemotiveerd raken door individuele uitdaging (Kuyper & Guldemon, 1997). Omdat de ADIT zich aanpast aan het niveau van de leerlingen en de vragen steeds uitdagender worden, zou dit de vwo-leerlingen mogelijk eerder uitdagen dan de havoleerlingen. Ook hebben havisten vaak een kleinere spanningsboog en zijn ze eerder afgeleid dan vwo-leerlingen. Voor de ADIT is het nodig om je anderhalf uur goed te kunnen concentreren. Het is mogelijk dat vwo'ers deze concentratie makkelijker op konden brengen dan de havisten en zo dus ook hoger konden scoren.

Omdat er een significant verschil is gevonden tussen havo en vwo-leerlingen in beide leerjaren is de totaalscore op de subtesten geen betrouwbare indicatie voor mogelijke inzetbaarheid van de ADIT in het eerste leerjaar. Dus ondanks dat de subtesten vlakken en woordbetekenis in totaal geschikt lijken te zijn om ingezet te worden in het eerste leerjaar, blijkt het verschil tussen havo en vwo-leerlingen significant te verschillen. Dit betekent dat deze subtesten mogelijk te weinig items bevatten om het niveau van vwo-leerlingen goed in te schatten. Om hier harde uitspraken over te doen moet er echter meer onderzoek verricht worden. Over het algemeen kan op basis van de resultaten beschreven in hoofdstuk 4 gesteld worden dat de huidige ADIT-itembank over onvoldoende items met een hoge moeilijkheidsparameter beschikt om ingezet te worden in leerjaar 1 van het voortgezet onderwijs.

5.2. In hoeverre bevat de huidige ADIT-itembank voldoende items met een hoge moeilijkheidsgraad ($B \geq 0$) om het intelligentieniveau van leerlingen uit het eerste leerjaar havo/vwo van het VO, nauwkeurig in te schatten?

Uit de resultaten van hoofdstuk 4 is te zien dat in totaal 14,5% van de leerlingen een IQ-score laat zien van ≥ 130 . Voor leerjaar 1 leerlingen is dit percentage 10,1% en voor leerjaar 2 leerlingen is dit percentage 24,6%. Alle drie de percentages

liggen boven de afgesproken grens van 2,5%. Er is in dit onderzoek een significant verschil gevonden tussen de IQ-uitkomsten van leerjaar 1 en leerjaar 2 leerlingen. Zoals hierboven al is benoemd kan dit verschil voortkomen uit het verschil in kennisopname tussen de beide leerjaren. Omdat er een significant verschil is gevonden, is de totale IQ-score geen betrouwbare indicatie. Er moet dus naar de uitkomsten per leerjaar worden gekeken om er een betrouwbare uitspraak over te kunnen doen. Zoals in hoofdstuk 3 is besproken is er in dit onderzoek alleen naar de IQ-uitslagen op de ADIT gekeken om achter het bereik van de moeilijkheidsparameter van de items te komen. In de literatuur is te vinden dat items met een hogere moeilijkheidsparameter correleren met een hogere IQ-score (Embretson & Reise, 2000; Lai, et al., 2011). Dit kan eveneens betekenen dat de frequentie en het bereik van de IQ-scores correleren met de frequentie en het bereik van de moeilijkheidsparameter van de aangeboden items. Omdat meer dan 2,5% van de leerlingen een TIQ-score heeft behaald van hoger dan 130, kan gesteld worden dat er sprake is van discrepantie tussen het niveau van de aangeboden vragen en het niveau van de havo en vwo-leerlingen in het eerste leerjaar van het voortgezet onderwijs.

Hierbij moet opnieuw vermeld worden dat deze resultaten voorzichtig geïnterpreteerd moeten worden aangezien er voor de beantwoording van deelvraag 2 gewerkt is met IQ-scores die voor groep 8 leerlingen zijn genormeerd. Met een nieuwe normering voor deze doelgroep kunnen andere scores worden verwacht. Dit zou ook kunnen verklaren waarom de IQ-scores voor leerjaar 1 leerlingen op de huidige ADIT hoger uitkomen dan men had verwacht.

5.3. Hoe wordt de afname van de ADIT door de leerlingen ervaren op de punten: lengte van de test, vraagstelling, vormgeving en de ervaren moeilijkheid?

Zo'n 65,4% van de leerlingen hebben de afname van de ADIT als positief ervaren. Over het algemeen zijn de meeste leerlingen tevreden over de afname op de onderdelen vormgeving, vraagstelling en de ervaren moeilijkheid. De leerlingen geven aan minder tevreden te zijn over de afnametijd van de test. Ze geven hierbij vooral aan dat ze niet altijd genoeg tijd hebben om een antwoord te geven op de gestelde vraag. Opvallend is dit wel aangezien de tijdslimiet vast is gesteld voor groep 8 leerlingen. Bij het construeren van de ADIT is de gemiddelde tijd die groep 8 leerlingen kwijt waren aan het beantwoorden van de vraag, plus 2 standaarddeviaties, gebruikt voor het vaststellen van de tijdslimiet per subtest. Hypothetisch zou er gesteld kunnen worden dat, aangezien leerlingen uit het eerste leerjaar van het VO over meer kennis beschikken, ze de vragen sneller zouden moeten kunnen beantwoorden. Een mogelijke verklaring voor dit verschil zou kunnen zitten in de moeilijkheid van de vragen. Uit een kleine analyse van de gegevens van groep 8 leerlingen bleek dat $\leq 1\%$ van de leerlingen de helft van de moeilijkste vragen aangeboden hebben gekregen. Het zou kunnen dat leerlingen meer tijd nodig hebben voor de moeilijkste vragen, aangezien deze meer beroep doen op hun kennis en vaardigheden dan de makkelijkere vragen.

Alleen op het onderdeel tijdsduur en op de algemene vraag of de leerlingen de afname als positief hebben ervaren, is er met minder dan 75% 'mee eens' gereageerd. Op basis van deze gegevens moet daarom geconcludeerd worden dat de afname over het algemeen niet als positief is ervaren. Dit moet echter met enige voorzichtigheid worden gesteld. Dit resultaat kan namelijk ontstaan omdat de leerlingen een test, met name een test waarvan ze geen uitslag krijgen, over het algemeen niet als positief of leuk ervaren. De deelvraag moet daarom vooral op subniveau worden bekeken: hier is namelijk de meeste informatie over de ervaring van de leerlingen met de ADIT te vinden. Op subniveau is te zien dat de leerlingen positief reageren op de verschillende onderdelen van de enquête: de vormgeving van de ADIT, de vraagstelling en de moeilijkheid van de ADIT. Over de lengte van de test zijn de leerlingen minder tevreden. Ze vinden de test te lang en gaven aan niet genoeg tijd te hebben om een vraag te beantwoorden.

5.4. Conclusie

De hoofdvraag van dit onderzoek luidde: In hoeverre is de huidige ADIT geschikt voor implementatie in het eerste leerjaar van het voortgezet onderwijs? Om deze vraag te kunnen beantwoorden is de huidige ADIT voorgelegd aan havo en vwo-leerlingen uit het eerste en tweede leerjaar van het voortgezet onderwijs. Uit de resultaten is af te leiden dat de huidige ADIT itembank over te weinig items beschikt om een nauwkeurige TIQ-meting te geven en dat er om deze reden een discrepantie bestaat tussen het niveau van de aangeboden vragen en het niveau van eerste jaars VO. Dit is met name van toepassing op leerlingen van vwo-niveau. Bij deze groep bleek namelijk de grootste discrepantie aanwezig te zijn, zowel in het eerste leerjaar als het tweede leerjaar.

Dit betekent dat de ADIT met zijn huidige itembank niet geschikt is voor implementatie in het eerste leerjaar. De huidige itembank zou uitgebreid moeten worden met items met een hogere moeilijkheidsparameter ($B > 2$) voordat de ADIT geschikt is voor implementatie in het eerste leerjaar. Voordat de implementatie kan plaatsvinden moeten de uitkomsten eerst nog genormeerd worden voor de nieuwe doelgroep. Pas dan kan de ADIT echt worden ingezet voor het eerste leerjaar VO. Daarnaast kwam uit de enquête dat de leerlingen over het algemeen de afname als positief hebben ervaren. Hier kwam echter ook uit dat leerlingen te weinig tijd hebben ervaren om goed antwoord te kunnen geven op de gestelde vragen. Daarnaast is het ook belangrijk om mee te nemen wat de leerlingen wel goed vonden. De leerlingen gaven namelijk aan dat ze de testinstructie helder en duidelijk vonden. Ze vonden de test uitdagend en goed passend bij hun niveau.

Als conclusie kan gesteld worden dat de ADIT met zijn huidige itembank niet geschikt is voor implementatie in het eerste leerjaar van het voortgezet onderwijs. In paragraaf 5.6 wordt verder ingegaan op de adviezen die gegeven kunnen worden voor het uiteindelijke doel: implementatie van de ADIT in het eerste leerjaar van het voortgezet onderwijs.

5.5. Discussie

Dit onderzoek heeft als doel een betrouwbaar en valide uitspraak te kunnen doen met betrekking tot de gehele onderzoekspopulatie. Om dit te realiseren is er gebruik gemaakt van een getrapte steekproef. Bij de opzet van dit onderzoek is er getracht een zo representatief mogelijke steekproef van alle leerjaar 1 leerlingen te verkrijgen. Tijdens de uitvoering van dit onderzoek is echter gebleken dat dit niet mogelijk was. De verdeling van leerlingen uit de vier regio's is in de steekproef niet representatief voor de verdeling die zich voordoet in de populatie. Eveneens is er een te kleine steekproef genomen om te kunnen spreken van generaliseerbaarheid van huidige resultaten naar de gehele onderzoekspopulatie. Ondanks dat dit bij de opzet van dit onderzoek wel getracht werd te bereiken, kan dit onderzoek wel gebruikt worden als eerste pilot of verkennend onderzoek. Hierbij moet nog aangegeven worden dat de groep vwo-leerlingen beduidend kleiner is dan de groep havo-leerlingen. Omdat een grotere steekproef niet mogelijk was, kan dit ten koste gaan van de betrouwbaarheid van de resultaten en uitspraken die betrekking hebben op deze groep. Ondanks dat er geprobeerd is de betrouwbaarheid op andere manieren te vergroten moet hier, in de interpretatie van de resultaten en conclusie, rekening mee worden gehouden. Waar verder rekening mee gehouden moet worden is dat de doelgroep die uiteindelijk voor dit onderzoek is gebruikt, niet de oorspronkelijke doelgroep was. Leerjaar 2 leerlingen zijn in dit onderzoek betrokken omdat ze het eind van het eerste leerjaar goed konden weerspiegelen. Aangezien het onderzoek twee maanden is uitgesteld, zijn deze leerlingen niet meer te vergelijken met eind leerjaar 1 leerlingen. Leerjaar 2 leerlingen zijn toch in dit onderzoek meegenomen omdat ze deze doelgroep alsnog het best kunnen weerspiegelen. Het kan echter dus zijn dat de resultaten anders waren geweest wanneer de ADIT in de oorspronkelijk geplande maand werd afgenomen.

In de loop van dit onderzoek moesten er aanpassingen gedaan worden om het onderzoek goed en haalbaar uit te kunnen voeren. Zo is er met een enkele school afgesproken dat de ADIT over meerdere lesuren afgenomen mocht worden. Er werd dus bijvoorbeeld het eerste lesuur een start gemaakt, waarna de leerlingen het tweede lesuur een andere les hadden en vervolgens het derde lesuur de ADIT alsnog gingen afronden. Hierdoor kan het zijn dat de leerlingen met een mindere concentratie de ADIT hebben afgerond dan dat ze begonnen. Dit kan ook van invloed zijn geweest op hun uitslagen. Ondanks deze uitzondering is er in dit onderzoek getracht de betrouwbaarheid zo goed mogelijk te waarborgen door het onderzoek zo transparant mogelijk uit te voeren en zijn er maatregelen genomen om de standaardisatie te verhogen. Zo is er in het onderzoek gebruikt gemaakt van betrouwbare en valide instrumenten om de hoofdvraag te kunnen beantwoorden. Zo werd er gebruik gemaakt van een COTAN voldoende beoordeelde test: de ADIT. De enquête laat daarnaast een voldoende onderlinge betrouwbaarheid zien (zie paragraaf 3.6).

Om ervoor te zorgen dat de ICT-omgeving bij elke school gelijk is, moet de school voor afname een 'ben ik geschikt'-test doen. Tijdens het onderzoek is dit echter niet gewaarborgd en sluit dit eventuele technische problemen op de dag zelf niet uit. Uit ervaring van de onderzoeker tijdens afnames is soms ook gebleken dat bij enkele testafnames sommige functies bleven vasthangen. Enkele leerlingen gaven hierover aan dat ze soms een antwoord niet konden geven omdat het programma vastliep en daardoor de tijd voor beantwoorden op was. Er moet rekening worden gehouden met mogelijke technische problemen bij het interpreteren van de resultaten en betrouwbaarheid van dit onderzoek. Scholen die aan het onderzoek meededen kregen dezelfde testinstructies toegestuurd, zodat de testinstructies voor elke leerling

hetzelfde zijn gebleven. Daarnaast worden de testinstructies op het beeldscherm van de leerlingen weergegeven. Deze testinstructie is voor elke leerling gelijk. Er moet in dit onderzoek rekening worden gehouden met het feit dat de afname van de ADIT niet gestandaardiseerd is. Dit houdt in dat de afname niet door dezelfde testleider afgenomen wordt en dat, ondanks dat is aangegeven de test af te nemen in een rustige omgeving, dit niet gewaarborgd kan worden. Dit kan tot gevolg hebben dat een klas bijvoorbeeld meer last heeft gehad van ruis tijdens de afname of dat een testleider meer bekwaam is in het beantwoorden van vragen van leerlingen. Deze factoren kunnen van invloed zijn op de betrouwbaarheid en validiteit van het onderzoek. Ook moet het gegeven dat leerlingen aangeven niet genoeg tijd te hebben om antwoord te geven op de vragen meegenomen worden in het analyseren van de testresultaten. Immers bestaat er een verschil tussen power-tests en speed-tests. Bij power-tests wordt er getest wat een leerling kan, bijvoorbeeld wat het rekenniveau van een leerling is. Bij een speed-test wordt er getest hoeveel rekenvragen een leerling binnen een bepaalde tijd correct kan beantwoorden. Beide soorten testen beogen dus een andere vaardigheid te meten. (Freeman, 1928) Belangrijk is om rekening te houden met het feit dat, ondanks dat de ADIT een power-test is, hij mogelijk gefungeerd kan hebben als een speed-test. Wanneer leerlingen langer de tijd hadden kunnen krijgen om over een antwoord na te denken, zouden ze hoger uit kunnen komen. De tijd kan ook van invloed zijn op het aantal goed gegeven antwoorden door het vergroten van het gokgedrag en de mogelijkheid tot het 'per ongeluk' goed of fout hebben van een vraag als gevolg van te weinig beantwoordingstijd. Het is goed om dit gegeven mee te nemen in de interpretatie van de conclusie.

Als laatste moet er bij de interpretatie van resultaten, conclusie en betrouwbaarheid rekening worden gehouden met eventuele menselijke reken- en invoerfouten. Alle resultaten zijn met de hand geselecteerd en ingevoerd. Hierdoor kunnen er eventuele fouten zijn gemaakt en hiermee kan zijn doorgerekend. De kans dat er fouten in de invoering zijn gemaakt is geprobeerd zo klein mogelijk te maken door elke berekening twee keer in en uit te voeren. Toch dient hier, bij de interpretatie van de resultaten en conclusie, rekening mee gehouden te worden, aangezien dit de betrouwbaarheid van de resultaten heeft vermindert.

5.6. Aanbevelingen

Vanuit de resultaten zijn aanbevelingen geformuleerd, gericht op het uiteindelijke doel van A-VISION: implementatie van de ADIT in het eerste leerjaar van het voortgezet onderwijs.

5.6.1 Aanbevelingen voor experimentele basis

Uiteindelijk zijn de hypothesen niet aangenomen. Mogelijke redenen hiervoor zouden kunnen liggen in de onderzoeksopzet. In deze paragraaf wordt er daarom beschreven wat er in vervolgonderzoek anders gedaan kan worden om deze verwachte effecten mogelijk wel te bereiken. Zoals in dit hoofdstuk al is verteld is dit onderzoek niet geheel representatief. Mogelijk is het aantal respondenten van invloed geweest op de resultaten en conclusie. Met name het aantal respondenten afkomstig vanuit het vwo. Deze groep is namelijk, ten opzichte van de gehele doelgroep, ondervertegenwoordigd. Hierdoor zijn de uitkomsten minder betrouwbaar voor deze groep leerlingen. Vervolgonderzoek zou zich daarom vooral ook op deze doelgroep moeten richten, aangezien uit het onderzoek komt dat deze doelgroep, in vergelijking met havoleerlingen, de moeilijkste vragen het vaakst aangeboden hebben gekregen. Daarnaast is, zoals hierboven al is benoemd, de oorspronkelijke datum voor afname verschoven, waardoor de keuze voor de huidige doelgroepen minder betrouwbaar is. Voor vervolgonderzoek kan geadviseerd worden om per maand of per kwartaal de ADIT voor te leggen en deze resultaten te analyseren. Op deze manier kan ook de generaliseerbaarheid van het onderzoek beter worden gewaarborgd.

Daarnaast is er tijdens dit onderzoek meerdere malen vanuit de scholen de tip gegeven om scholen eerder te benaderen mee te doen met dit onderzoek. Bij voorkeur hebben scholen het liefst dat ze benaderd worden in de maand mei of juni, omdat ze dan de nieuwe jaarplanning voor de school gaan maken en met een groot onderzoek zoals dit rekening kunnen houden. Tijdens dit onderzoek is namelijk gebleken dat het best lastig was om scholen bereid te krijgen om mee te doen aan een onderzoek waar ze zelf best nog wat tijd in moeten steken. Dit verklaart ook waarom de gewenste steekproefgrootte in dit onderzoek niet is behaald. Voor vervolgonderzoek zou daarom geadviseerd kunnen worden om scholen in deze maanden te benaderen. Hierdoor wordt de kans vergroot dat ze mee kunnen/willen doen. Hiermee wordt de steekproef groter en vergroot dit de betrouwbaarheid van het onderzoek.

5.6.2 Aanbevelingen voor vervolgstappen

Het uiteindelijke doel is om de ADIT geïmplementeerd te krijgen in het eerste leerjaar van het voortgezet onderwijs. Op basis van dit onderzoek blijkt dat dit nu, met de huidige itembank, nog niet kan. Hieronder worden mogelijke vervolgstappen besproken om toch tot het doel te komen.

Als eerste vervolgstap is het natuurlijk belangrijk om de huidige itembank, voor alle subtesten, uit te breiden met vragen met een hogere moeilijkheidsparameter ($B > 2$). Dit moet ervoor zorgen dat er, met name voor vwo-leerlingen, genoeg vragen in de itembank zitten om hun niveau met een hoge nauwkeurigheid in te kunnen schatten. Leraren en leerlingen kunnen helpen in het meedenken van het soort vragen dat bijgemaakt zouden kunnen of moeten worden. Dit verschilt immers per subtest. Bij rekenopgaven zouden de toe te voegen items bijvoorbeeld makkelijker te bepalen zijn dan bij taalonderdelen. Aan docenten zou gevraagd kunnen worden naar vragen waarvan zij denken dat ze moeilijk genoeg zijn voor vwo-leerlingen. Hierop gebaseerd zouden vragen ontwikkeld kunnen worden. Eventueel zouden enkele voorbeeldvragen aan vwo-leerlingen voorgelegd kunnen worden, met de vraag op een schaal van 1 tot 5, hoe moeilijk/uitdagend ze de vraag vinden. Op basis van deze informatie zou doorontwikkeld kunnen worden en zouden er meer vragen ontwikkeld kunnen worden. Dit advies wordt gegeven omdat het jammer zou zijn wanneer de itembank aangevuld zou worden, maar uiteindelijk toch blijkt dat de items die aangevuld zijn nog niet uitdagend genoeg zijn. Wanneer nieuwe items aan de itembank zijn toegevoegd is het daarnaast belangrijk om deze itembank weer voor te leggen aan de doelgroep, om te zien of er, met de uitbreiding, genoeg items in de vernieuwde itembank zitten om nauwkeurig het niveau te kunnen bepalen. Aanbevolen kan worden om vervolgonderzoek specifiek te richten op deze hogere klassen, de vwo-groepen van het eerste leerjaar.

Daarnaast zou geadviseerd kunnen worden om te kijken naar het tijdsaspect van de ADIT. Omdat toch veel leerlingen aangeven dat ze te weinig tijd hebben bij het beantwoorden van de items, is het aan te raden hier een vervolgonderzoek naar te doen. Mogelijk kan het zijn dat het tijdsaspect van invloed is geweest op de uitkomsten van dit onderzoek. Mogelijk kan het ook zo zijn dat het tijdsaspect verschilt per doelgroep. Dus wat groep 8 leerlingen gemiddeld kwijt zijn voor het beantwoorden van een vraag, kan verschillen van wat eerstejaars leerlingen gemiddeld kwijt zijn voor het beantwoorden van dezelfde vraag. Omdat hier verder nog geen informatie over verzameld is en omdat leerlingen dit toch veel aangaven, is het aan te raden hier actie op te ondernemen in vervolgonderzoek. Men kan overwegen om dit mee te nemen in het vervolg normeringsonderzoek. Wanneer er namelijk items worden toegevoegd aan de huidige itembank moet de ADIT eerst worden genormeerd door middel van een normeringsonderzoek. Hierna kan deze pas worden geïmplementeerd. Tijdens het normeringsonderzoek zou dan ook per gemaakte item (*server-sided*) gemeten kunnen worden hoe lang er over de beantwoording van ieder item gedaan werd. Van iedere subtest kan vervolgens de gemiddelde beantwoordingstijd en de standaarddeviatie berekend worden. Vervolgens zou voor deze doelgroep dezelfde overweging gemaakt kunnen worden bij het instellen van de tijdslimiet. Dit betekent: de gemiddelde tijd per item plus twee standaarddeviaties, af te ronden op 5 hele seconden.

Een andere mogelijkheid is om de leerlingen langer de tijd te geven bij de moeilijkere items, zoals ze dit bij andere intelligentietesten ook doen. Voordat hier veranderingen in aangebracht worden, is het belangrijk om te weten of de leerlingen benadeeld worden door de tijdslimiet. Onderzoek hiernaar is dus noodzakelijk om tot een goed antwoord en adviesplan te komen.

5.7. Afsluiting

Ter afsluiting wordt er hier een persoonlijke notitie neergezet. Na wat gesprekken met leerkrachten en leerlingen is gebleken wat het effect is van het afnemen van een intelligentietest bij leerlingen. Leerlingen zijn erg benieuwd naar hun eigen intelligentieniveau en kunnen zich hierdoor uitgedaagd voelen om zich te verbeteren en zo een hogere score te behalen. Leerlingen kunnen het ook eng vinden en kunnen nerveus zijn voor de uitkomst van een intelligentietest. Hier is gelijk ook de kracht van het adaptieve testen terug te vinden. Na met leerlingen gesproken te hebben, viel het voor de meeste wel mee. Het was niet zo eng als ze hadden verwacht en de vragen konden ze allemaal wel beantwoorden. Ook leraren gaven dit aan en vonden het prettig om te zien hoe geconcentreerd en goed hun leerlingen de test aan het maken waren. Voor de leerlingen was het niet altijd even prettig en leuk om te doen, maar ik heb positieve en leuke reacties teruggekregen die mee te nemen zijn in het vervolgetraject van de ADIT eerste leerjaar VO.

Literatuurlijst

- Alem, van, V., & Horst, ter, S. (2013). *Kennisnet: Digitaal toetsen in de basisschool*. Nederland, Leiden: OBT de Brink
- Alfonso, V.C., Flanagan, D.P., & Radwan, S. (2005). The impact of the Cattell-Horn-Carroll theory on test development and interpretation of cognitive and academic abilities. In D.P. Flanagan & P.L. Harrison ((Eds.), *Contemporary intellectual assessment: Theories, test and issues*(2nd ed., pp. 185-202). Geraadpleegd op <http://faculty.mwsu.edu/psychology/dave.carlston/IQ/alfonso.pdf>
- A-VISION. (n.d.). Over ons. Geraadpleegd op <http://a-vision.nu/pages/Over+ons.html#autoscroll>
- Carson, K., Gillon, G., & Boustead, T. (2011). Computer-administrated versus paper-based assessment of school-entry phonological awareness ability. *Asia Pacific Journal of Speech, Language and Hearing*, 14(2), 85-101.
- Centraal Bureau voor de Statistiek. (2016). VO; leerlingen, leeftijd. Geraadpleegd op <http://cbs.overheidsdata.nl/80041ned>
- Centraal Bureau voor de Statistiek. (2016). VO; leerlingen, regio in detail. Geraadpleegd op <http://cbs.overheidsdata.nl/80042ned>
- CITO. (n.d.). Adaptieve toetsen. Geraadpleegd op http://www.cito.nl/onderwijs/primair%20onderwijs/lvs_toetsen/digitaal_toetsen/adaptief
- Coubergs, C., Struyven, K., Engels, N., Cools, W., & de Martelaer, K. (2013). *Binnenklasdifferentiatie: Leerkans voor alle leerlingen*. Leuven, Nederland: Acco
- Daamen, W. (2015). *Wat werkt bij het implementeren van jeugdinterventies?* Utrecht, Nederland: Nederlands Jeugdinstituut
- Dijk, van, H. (2012). NIO Nederlandse Intelligentietest voor Onderwijsniveau. Beter Begeleiden Digitaal. Online publicatie. Lbbo.nl
- Dijk, van, H., & Tellegen, P. (2004). *NIO Nederlandse Intelligentietest voor Onderwijsniveau: handleiding en verantwoording*. Amsterdam, Nederland: Boom test
- Drenth, P., & Sijsma, K. (2012). *Testtheorie: Inleiding in de theorie van de psychologische test en zijn toepassingen* (Herz. ed.). Houten, Nederland: Bohn Stafleu van Loghum
- Embretson, S. E., & Reise, S. P. (2000). *Item Response Theory for Psychologists*. Mahwah, New Jersey: Lawrence Erlbaum Associates
- Evers, A., Vliet-Mulder, van, J. C. & Groot, C. J. (2000). *Documentatie van tests en testresearch in Nederland, deel I en II*. Assen: Van Gorcum.
- Freeman, F. S. (1928). Power and speed: their influence upon intelligence tests scores. *Journal of Applied Psychology*, 12(6), 631-635
- Groninger Intelligentietest Voortgezet Onderwijs (GIVO). (n.d.). Geraadpleegd op http://www.toetswijzer.nl/toetsinfo.asp?Mode=Detail&toe_id=74
- Houtveen, A. A. M., & Reezigt, G. J. (2000). *Succesvol adaptief onderwijs*. Alphen a/d Rijn: Samsom

- Joosten-ten Brinke, D., Weber, A., & Saveski, G.L. (2016). *Digitaal grootschalig summatief toetsen in het primair onderwijs: onderzoek rapportage in opdracht van kennisnet*. Geraadpleegd op https://www.kennisnet.nl/fileadmin/kennisnet/leren_ict/digitaal_toetsen/bijlagen/Digitaal_grootschalig_summatief_toetsen_in_het_primair_onderwijs.pdf
- Kuyper, H., & Guldemon, H. (1997). *Studievaardigheden en huiswerkgedrag in de bovenbouw van HAVO en VWO*. Groningen, Nederland: GION.
- Kuyper, H., van der Werf, M. P. C., & Lubbers, M. J. (1999). *Tussen basisvorming en studiehuis: de bovenbouwstudie van VOCL'93*. Groningen, Nederland: GION.
- Lai, J. S., Cella, D., Choi, S., Junghaenel, D.H., Christodoulou, C., Gershon, R., & Stone, A. (2011). How item banks and their application can influence measurement practice in rehabilitation medicine: a PROMIS fatigue item bank example. *Archives of physical medicine and rehabilitation*, 92(10), S20-S27
- Lycaeus. (n.d.). Lycaeus Juridisch Woordenboek. Geraadpleegd op <http://www.juridischwoordenboek.nl/woordenboekdis.html#13202>
- Van Minden, J.J.R. (2010). *Alles over psychologische test* (26e Herz. ed.). Amsterdam, Nederland: Business Contact
- Montgomery, J. M., & Cutler, J. (2013). Computerized Adaptive Testing for Public Opinion Surveys. *Political Analysis*, 21(2), 172-192. [Hhps://doi.org/10.1093/pan/mps060](https://doi.org/10.1093/pan/mps060)
- Padmos, B. (2001). *De Scriptie supporter: Een doeltreffende aanpak van je scriptie*. Apeldoorn, Nederland: Garant
- Pameijer, N. (2011). Waarom een ontwikkelingsperspectief meer is dan IQ en leerrendement. *Tijdschrift voor Orthopedagogiek: Onderzoek en Praktijk*, 50, 461-472
- Resing, W., & Drenth, P. (2007). *Intelligentie: weten en meten* (2nd ed.). Amsterdam, Nederland: Nieuwezijds
- Schoonman, W. (2011). *Methodologische achtergronden van het testinstrumentarium*. Geraadpleegd op <http://www.hro-businesspartners.com/images/Downloads/overzicht-methodologische-achtergronden-testinstrumentarium.pdf>
- Sectorakkoord VO 2014-2017. (2014)/ Geraadpleegd op http://www.schoolaanzet.nl/fileadmin/contentelementen/school_aan_zet/Sectorakkoord-VO-OCW.pdf
- Tellegen, P. (2014). De COTAN ter discussie. Geraadpleegd op http://scholar.googleusercontent.com/scholar?q=cache:woSIFGzKJ1cJ:scholar.google.com/+cotan+steekproefgrootte&hl=nl&as_sdt=0,5
- Tomlinson, C. A., & Imbeau, M. B. (2010). *Leading and Managing a Differentiated Classroom*. Alexandria, VA: ASCD
- Universiteit van Amsterdam (2010). @daptief toetsen. Geraadpleegd op <http://www.uva.nl/binaries/content/documents/personalpages/k/l/s.klinkenberg/en/tab-two/tab-two/cpitem%5B11%5D/asset?1355373527110>
- Valk, G. (2005). 'School overschat waarde IQ-test': Intelligentietesten voor kinderen zijn onbetrouwbaar, zegt de maker van zo'n test. NRC, pp 1, 3. Geraadpleegd op <https://www.nrc.nl/>
- Vegt, E. R., & Metselaar, E. E. (2013). Verantwoording van de ADIT: Een adaptieve intelligentietest voor groep 8 van het basisonderwijs. Geraadpleegd via A-VISION
- Verhoeven, N. (2011). *Wat is onderzoek? Praktijkboek methoden en technieken voor het hoger onderwijs* (4rd ed.). Den Haag, Nederland: Boom Lemma

Weiss, D. J. (1985). Adaptive testing by computer. *Journal of Consulting and Clinical Psychology*, 53, 774-789

Weiss, D. J., & Kingsbury, G. (1984). Application of computerized adaptive testing to educational problems. *Journal of Educational Measurement*, 21(4), 361-375.

Zand, van, D., Pelt, D., & Schrijver, M. (2015). Handleiding Adaptieve Capaciteiten Test Algemene Intelligentie: Versie 1.0. Geraadpleegd op <http://www.test-toolkit.nl/wp-content/uploads/2015/10/HandleidingACT-Versie-1.0-Final.pdf>

Zee, van der, F. (2005). Het onderscheid tussen kwalitatief en kwantitatief onderzoek is gradueel. KWALON, 2. Geraadpleegd op http://www.boomlemmatijdschriften.nl/tijdschrift/KWALON/2005/2/KWALON_2005_010_002_00

Zeeuw, de, J., Dekker, R., & Resing, W. C. M. (2004). *Algemene Psychodiagnostiek 1. Testmethoden*. Geheel herziende druk. Leiden: PITS

Bijlagen

B1. Overzicht intelligentietesten en rol ADIT

Voor het middelbaar onderwijs bestaan er meerdere COTAN gecertificeerde intelligentietesten. Deze zijn met hun kenmerken in tabel 12 weergegeven. In het rood is de ADIT weergegeven als mogelijke plek die deze intelligentietest, wanneer geïmplementeerd, in kan nemen tussen de al bestaande tests (Evers, van VlietMulder, & Groot, 2000⁽¹⁾; Groninger Intelligentietest Voortgezet Onderwijs (GIVO)⁽²⁾, n.d.; van Dijk & Tellegen, 2004⁽³⁾; Vegt & Metselaar, 20013⁽⁴⁾).

Tabel 12 <i>Vergelijking intelligentietesten doelgroep</i>					
	WISC-III⁽¹⁾	Son-R⁽¹⁾	GIVO⁽²⁾	NIO⁽³⁾	ADIT⁽⁴⁾
Leeftijd	6-16 jr.	5,5-16 jr.	12-16 jr.	12-16 jr.	Groep 8
Meet pretentie	Algemene intelligentietest	Niet-verbale intelligentietest	t.b.v. advisering m.b.t. type VO.	Algemene intelligentietest	Algemene intelligentietest
Aantal subtests	13 subtests	7 subtests	7 subtests	6 subtests	7 subtests
Afnameduur	75 minuten	90 minuten	135 minuten	≈ 120 minuten	90 minuten
Voor hoeveel?	Individueel	Individueel	Kleine groepen/ individueel	Individueel en groepsgewijs	Individueel en groepsgewijs
Afnamevorm	Pen en papier & digitaal	Pen en papier	Pen en papier	Pen en papier	Digitaal
Scoring	3 IQ-scores 3 factorscores	1 IQ-score	Verbale intelligentie, rekenkundig niveau, rekenkundige vaardigheden en ruimtelijk inzicht	Verbale-, Symbolische-, totale intelligentie index. En schooltype.	Non-verbaal, verbaal en totale intelligentie.
COTAN-Beoordeling	Betrouwbaarheid: V Begripsvaliditeit: V Criteriumvaliditeit: O Normen: V	Betrouwbaarheid: G Begripsvaliditeit: G Criteriumvaliditeit: G Normen: G	Betrouwbaarheid: G Begripsvaliditeit: G Criteriumvaliditeit: G Normen: V	Betrouwbaarheid: G Begripsvaliditeit: G Criteriumvaliditeit: G Normen: G	Betrouwbaarheid: G Begripsvaliditeit: V Criteriumvaliditeit: V Normen: V
Afname door	Psychologen, orthopedagogen of gekwalificeerde testassistenten	Psychologen, orthopedagogen of gekwalificeerde testassistenten	Psychologen, orthopedagogen of gekwalificeerde testassistenten	Psychologen, orthopedagogen of gekwalificeerde testassistenten	Iedereen die bekend is met de test

Opmerking: Bronnen gebruikt voor het samenstellen van de tabel staan boven de tabel weergegeven met bijbehorende nummer

B2. Biografische informatie populatie

Geslacht:

Tabel 13 Onderzoekspopulatie naar geslacht		
	Aantallen	Percentage
Jongens	108.469	51.3%
Meisjes	103.119	48.7%
Totaal	211.593	100%

Herkomst:

Tabel 15 Onderzoekspopulatie naar herkomst		
	Aantallen	Percentage
Autochtoon	156.982	74.2%
Allochtoon	51.663	24.4%
Westerse Allochtoon	14.362	6.8%
Niet-westerse Allochtoon	37.301	17.6%
Onbekend	2.948	1.4%
Totaal	211.593	100%

Leeftijd:

Tabel 14 Onderzoekspopulatie naar leeftijd		
	Aantallen	Percentage
≤ 11 jaar	3.661	1.7%
12 jaar	129.826	61.4%
13 jaar	66.408	31.4%
14 jaar	5.906	2.8%
15 jaar	1.714	0.8%
≥ 16 jaar	4.078	1.9%
Totaal	211.593	100%

Leerlingen per provincie:

Tabel 16 Onderzoekspopulatie naar onderwijsprovincie		
	Aantallen	Percentage
Groningen	7.458	3.5%
Friesland	8.972	4.2%
Drenthe	5.976	2.8%
Noord totaal	22.406	10.5%
Overijssel	16.095	7.6%
Flevoland	6.058	2.9%
Gelderland	26.305	12.4%
Oost totaal	48.458	22.9%
Utrecht	16.205	7.7%
Noord-Holland	32.786	15.5%
Zuid-Holland	44.377	21%
Zeeland	4.344	2.1%
West totaal	97.712	46.3%
Noord-Brabant	31.566	14.9%
Limburg	11.451	5.4%
Zuid totaal	43.017	20.3%
Totaal	211.593	100%

B3. Standaard e-mail contact scholen



Beste meneer/mevrouw,

Ik ben Kyra van Rinsum en ik studeer Toegepaste Psychologie in Deventer. Ik doe mijn afstudeeronderzoek bij A-VISION. A-VISION is een landelijk werkend onderwijsadviesbureau. Ze zijn gespecialiseerd in het ontwikkelen en afnemen van adaptieve, digitale tests, ten behoeve van de juiste plaatsing in het onderwijs. Zo hebben ze ook de ADIT (adaptieve digitale intelligentietest) ontwikkeld.

De ADIT kan op dit moment worden afgenomen in groep 8. Wij merken dat er veel vraag is naar een intelligentietest die zich afstemt op het niveau van de leerling in het eerste leerjaar van het voortgezet onderwijs, daarom willen we de ADIT gaan normering voor het eerste leerjaar van het VO. Om de ADIT aan te kunnen bieden is er echter onderzoek nodig. Voor mijn afstudeerproject mag ik gaan onderzoeken of de ADIT geschikt is voor het eerste leerjaar.

Wat de ADIT uniek maakt, in vergelijking met andere intelligentietesten voor deze doelgroep (NIO en de WISC-III) is het adaptieve karakter. Dit betekent dat de test zich aanpast aan het niveau van de leerling. Daarnaast zou de ADIT groepsgewijs en individueel afgenomen kunnen worden, zodat meerdere leerlingen hiervan gebruik kunnen maken. Voor meer informatie verwijs ik u graag naar bijgevoegde informatiebrief. Belangrijk om te weten is dat de ADIT niet in zijn geheel gelijk afgenomen hoeft te worden.

Aangezien er geen uitslag gegenereerd kan worden op basis van dit onderzoek en ik het erg waardeer als u mee zou willen werken, kunnen de deelnemende leerlingen kosteloos onze LEV-L (Leervoorkeurenlijst) of de MITT ((Meervoudige Intelligentie/ Talenten Test) maken. Hiervan ontvangt u per leerling een persoonlijk profiel. Een voorbeeld is bijgevoegd.

Het normeringsonderzoek kan afgenomen worden **tot eind Januari 2017**. Ik kan alle hulp om het tijdig af te ronden goed gebruiken. Voor dit onderzoek ben ik op zoek naar **havo** en **vwo**-leerlingen uit het **eerste en tweede leerjaar**, die mij kunnen helpen.

Mocht u mee willen doen met uw leerlingen, kunt u zich aanmelden via KvanRinsum@a-vision.nu. Wij zullen u dan digitaal het aanmeldformulier toesturen.

Mocht u nog vragen hebben, stel ze gerust!
Graag zie ik uw reactie tegemoet. Alvast hartelijk dank.

Vriendelijke groet,
Kyra van Rinsum
Afstudeerster Toegepaste Psychologie

Aanwezig: maandag t/m donderdag van 8:30 - 16:30
Tel: 055-5400333

Adres Laan van Westenenk 162 7336 AV Apeldoorn	Telefoon/fax 055 - 540 03 33 055 - 202 09 05		E-mail/website info@a-vision.nu www.a-vision.nu	Twitter/Linkedin/Facebook @AVISIONTesten a-vision-onderwijsadviesbureau facebook.com/AVISIONTesten
--	--	---	--	---

B4. Testinstructie ADIT

Algemene instructie

Om het testen goed te laten verlopen is het nodig dat u zelf vooraf deze instructie goed doorneemt en de inlogprocedure en de test doorloopt.

Er kan worden ingelogd zoals hieronder beschreven. Om als begeleider de tests te kunnen doorlopen is er een inlogcode voor de testbegeleider aangemaakt (zie overzicht inlogcodes).

De inlogprocedure door de leerlingen moet nauwkeurig gebeuren. Gebruikersnamen en wachtwoorden staan op het leerlingblad dat u hebt gekregen. De inloggegevens van alle leerlingen staan vermeld op het groepsoverzicht.

Naar het testprogramma van A-VISION

- Ga naar de website waar de tests staan door de Internet Explorer-browser te starten en in de adresbalk te typen: <http://www.toetsendigitaal.nl>
- U komt op een inlogscherm waar de leerlingen met hun gebruikersnaam en wachtwoord moeten inloggen. Daarmee krijgen ze toegang tot de test.

Naar een test

- Op het leerlingblad staan de naam van de leerling, de gebruikersnaam en het wachtwoord om toegang te krijgen tot de ADIT.
- Een gebruikersnaam bestaat uit 2 cijfers, 2 HOOFDLETTERS en 3 cijfers (bijv.: 12NA123).
- Het wachtwoord bestaat uit 2 HOOFDLETTERS en 3 cijfers (bijv. NA123).
- Zorg er voor dat bij elke computer een leerlingblad klaar ligt voordat de leerlingen het lokaal binnenkomen en dat de leerlingen bij de computer plaatsnemen waar het leerlingblad met hun naam ligt.
- Vertel de leerlingen dat ze niet aan het toetsenbord mogen komen voordat dat gezegd wordt.

De instructie voor de leerlingen om in te loggen:

- a. Typ achter 'gebruikersnaam' de cijfers en letters in die op het leerlingblad achter 'gebruikersnaam' staan. Doe dat nauwkeurig en denk erom de letters zijn HOOFDLETTERS. Controleer of je het goed hebt ingetoetst.
- b. Typ achter 'wachtwoord' het wachtwoord in dat op het leerlingblad staat. Doe dat nauwkeurig en denk erom de letters zijn HOOFDLETTERS. Controleer het goed.
- c. Klik op [Log in].
- d. Je krijgt nu een pagina te zien met daarop je eigen naam en daaronder 'overzicht testen'
- e. Klik op 'overzicht testen'
- f. Klik nu bij de ADIT op 'start de test'.

Wanneer het eerste scherm van de test is geopend, kan de algemene instructie met de leerlingen gezamenlijk worden doorgenomen.

Leerlingen moeten een antwoord geven voor ze naar de volgende vraag kunnen, tenzij de tijd die voor een vraag staat verstrijkt.

De test bestaat uit zeven onderdelen. Alle onderdelen beginnen met oefeningen, daarna wordt gestart met het onderdeel. Als een onderdeel gemaakt is krijgt het de status 'klaar' en kan de test niet opnieuw worden gemaakt.

Loopt de computer vast bij een bepaalde vraag, klik dan op de knop 'vernieuwen' (vakje met 2 pijlen rechts van het inlogvak op de werkbalk) of op F5. Meestal komt dan de goede pagina weer in beeld. Niet op Enter of Delete klikken!

Als het binnenhalen van de vragen traag verloopt of computers vastlopen, kan dat komen doordat het gebruik van de Smartboards de internetverbinding 'verstopt.' Het is beter om dan een ander moment te kiezen waarop de internetverbinding minder bezet is.

Het is de bedoeling dat leerlingen **alle onderdelen van de test achter elkaar maken** zonder uit te loggen. Er kan niet worden uitgelogd tijdens een testonderdeel, wel als een onderdeel is afgesloten.

Mocht er zich onverhoopt een storing voordoen dan kunt u contact opnemen met uw systeembeheerder of anders met het kantoor van A-VISION (055 5400333).

Heel veel succes!

B5. 'Ben ik geschikt'-stappenplan

Voordat uw leerlingen een toets (of meerdere toetsen) gaan maken, willen wij graag inzicht krijgen in uw ICT-omgeving. Wij willen kunnen zien of uw ICT-omgeving geschikt is voor de afname van het onderzoek om problemen te voorkomen tijdens de testafname. Daarnaast kunnen wij, indien nodig, eraan bijdragen uw ICT-omgeving geschikt te maken. Hiervoor dient u gebruik te maken van ons systeem: 'ben ik geschikt'. Dit kan tussen 9.00 en 14.00 uur (bij voorkeur niet tijdens de schoolpauzes). U doorloopt hiervoor onderstaande stappen.

1.

Neem plaats achter een van de computers die gebruikt wordt bij de afname van de toets.

LET OP: Het is van belang dat dit een computer is die door de leerlingen tijdens de toetsafname wordt gebruikt. Een computer van een leerkracht of ICT'er heeft vaak meerdere updates e.d. gehad.

LET OP: U dient de check uit te voeren op elk soort computer die gebruikt wordt voor de testafname alsmede bij verschillende besturingssystemen. Gebruiken de leerlingen allemaal dezelfde soort computer/laptop/tablet, dan hoeft de check slechts op één apparaat uitgevoerd te worden. Gebruikt de ene leerling een tablet en de andere leerling een laptop, voer de check dan op beide soorten apparaten uit.

2.

Ga naar: www.routewijs.nl/benikgeschikt.
Hiernaast gegeven scherm verschijnt. Vul uw gegevens in en druk daarna op "eerste controle uitvoeren".

LET OP: Bij het inloggen wordt gevraagd om uw groepscode. Wanneer u deze nog niet heeft ontvangen mag u hier 000 invullen.

IS UW APPARAAT GESCHIKT VOOR ONZE TESTEN?

Let op: Tijdens het testen van uw apparaat zullen er video en audiofragmenten worden afgespeeld. Voor de audiofragmenten dient u luidsprekers of een hoofdtelefoon aangesloten te hebben.

VUL UW GEGEVENS IN

Naam
Functie
School / SWV
Plaats
ICT Dienstverlener
E-Mailadres
Telefoonnummer
Testgroep code
Eerste controle uitvoeren

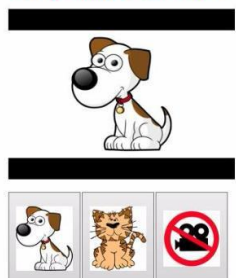
© 2015 - A-VISION B.V.

3.

Er wordt een controle uitgevoerd en vervolgens verschijnt er een groene balk. Klik vervolgens op "volgende stap".
Er wordt nu gecontroleerd of uw computer de plaatjes goed weer kan geven. Klik aan welke afbeelding u ziet.

IS UW APPARAAT GESCHIKT VOOR ONZE TESTEN?

GEEFT U AAN WELKE AFBEELDING U ZIET.



© 2015 - A-VISION B.V.

4.

Wanneer u bovenstaande stap heeft uitgevoerd wordt er gecontroleerd of uw computer geluidsfragmenten goed af kan spelen. Klik aan of u het audio fragment wel of niet kunt horen.

IS UW APPARAAT GESCHIKT VOOR ONZE TESTEN?

ER WORDT EEN AUDIOFRAGMENT AFGESPEELD. KUNT U DIT HOREN?



JA	NEE
----	-----

© 2015 - A-VISION B.V.

5.

Indien u ICT-omgeving geschikt is verschijnt hiernaast weergegeven scherm.

Indien uw ICT-omgeving niet geschikt is, kan bijvoorbeeld hiernaast weergegeven scherm verschijnen. Als u de melding krijgt dat er een probleem is en u kunt dit niet zelf oplossen, neem dan contact op via 055 5400333

IS UW APPARAAT GESCHIKT VOOR ONZE TESTEN?



UW APPARAAT IS GESCHIKT VOOR ONZE TESTEN.

© 2015 - A-VISION B.V.

IS UW APPARAAT GESCHIKT VOOR ONZE TESTEN?

De snelheid van uw verbinding is te laag. Hierdoor kunnen wij niet garanderen dat de test goed verloopt. Wij adviseren u om deze verbinding op te waarderen of om de bevestigingen verspreid over de dag te testen.

© 2015 - A-VISION B.V.

Bedankt voor uw medewerking!

Team A-VISION

B6. Enquête

Wij zouden graag willen weten hoe jij de ADIT hebt ervaren. Met jouw mening kunnen we de ADIT namelijk nog beter maken. In onderstaande stellingen wordt daarom om jouw mening gevraagd. Hiervoor zijn 4 antwoordmogelijkheden (van niet mee eens tot helemaal mee eens). Een fout antwoord kun je niet geven, het gaat immers om jou mening! De antwoorden worden vertrouwelijk behandeld. Het invullen van de vragenlijst zal enkele minuten duren.

Dankje wel!

	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>
1. De testafname van de ADIT vond ik positief	☐	☐	☐	☐
2. Het feit dat de test op een computer wordt afgenomen vind ik positief	☐	☐	☐	☐
3. Het feit dat de test adaptief is vind ik positief	☐	☐	☐	☐
Lengte van de test				
4. De lengte van de test vond ik goed	☐	☐	☐	☐
5. Ik kon mijn concentratie aan het begin van de test er goed bijhouden	☐	☐	☐	☐
6. Aan het eind van de test kon ik mijn concentratie er ook nog goed bijhouden	☐	☐	☐	☐
7. Ik had voldoende tijd om een vraag binnen de tijd in te vullen	☐	☐	☐	☐
Vraagstelling				
8. Ik vond de instructie voorafgaand aan de vragen helder en duidelijk	☐	☐	☐	☐
9. Ik vond de vragen helder en duidelijk	☐	☐	☐	☐
10. De woordkeuze in de test vond ik duidelijk en goed	☐	☐	☐	☐
11. De manier waarop de vragen gesteld worden vond ik duidelijk en goed	☐	☐	☐	☐
Vormgeving				
12. De plaatjes gebruikt in de test vond ik helder en duidelijk	☐	☐	☐	☐
13. De opmaak van de testpagina (de webpagina) vond ik duidelijk	☐	☐	☐	☐
Moeilijkheid				
14. Was de test voldoende moeilijk?	☐ Ja, hij was voldoende moeilijk ☐ Nee, de test was te makkelijk ☐ Nee, de test was te moeilijk			
15. Heeft de test je uitgedaagd tot het hoogste van je niveau?	☐ Ja, hij was voldoende uitdagend ☐ Nee, de test was te veel uitdagend ☐ Nee, de test was te weinig uitdagend			

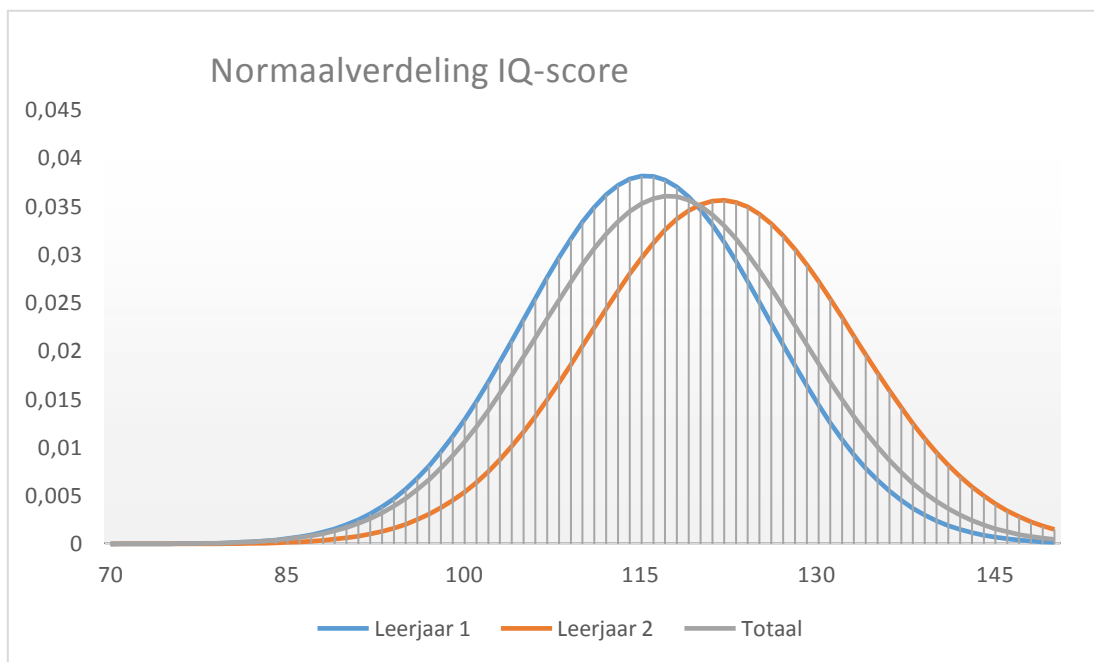
Wat zijn jouw suggesties voor verbetering van de ADIT?

B7. Overzicht afnameprocedure ADIT

1.	a. Aanmelding De school verkrijgt het aanmeldformulier (een Excel-bestand) voor de testafname van de ADIT per e-mail naar aanleiding van telefonisch contact met A-VISION. b. Invullen van het aanmeldformulier De school vult het aanmeldformulier als volgt in: - op blad 1 van het aanmeldformulier komen de schoolgegevens te staan - op blad 2 worden de leerlinggegevens (waaronder naam, geboortedatum en leerjaar) ingevuld. c. Retourneren van het aanmeldformulier De school verzendt het ingevulde aanmeldformulier naar de contactpersoon bij A-VISION
2.	Verwerking aanmeldgegevens Na ontvangst van het aanmeldformulier wordt door A-VISION de ADIT gekoppeld aan de testgroep. Elke leerling krijgt een unieke gebruikersnaam en wachtwoord.
3.	Wijze van afname Met de school wordt overlegd over het afnemen van de ADIT via de schoolserver of via de server van A-VISION. a. Schoolserver: met de ICT-er worden afspraken gemaakt over plaatsing van het testprogramma op de schoolserver. De instructie en verdere informatie worden per e-mail verstuurd. b. Server van A-VISION: het testprogramma staat klaar op de server van A-VISION. De leerlingen kunnen rechtstreeks inloggen via www.routewijs.nl
4.	Testafname A-VISION stuurt de inloggegevens van de ADIT (overzicht van de inlogcodes per groep en individuele leerlingbladen) naar de school. De leerlingbladen worden bij de computers gelegd om te gebruiken bij de testafname. De testafname wordt in een rustige testruimte en onder deskundige leiding afgenomen. Dit kan een werknemer van A-VISION of van de school zijn. A-VISION is altijd bereikbaar voor vragen en of opmerkingen en om hulp te bieden tijdens de afname.
5.	Verwerking en rapportage Zodra alle leerlingen de ADIT gemaakt hebben, krijgt A-VISION dit door en worden de testresultaten verwerkt. Voor elke leerling wordt een rapportage aangemaakt en voor de school beschikbaar gesteld.

B8. Normaalverdeling TIQ

In onderstaande figuur (3) zijn de normaalverdelingen weergegeven, zoals deze in dit onderzoek gevonden zijn voor de totale steekproef, voor leerjaar 1 leerlingen uit de steekproef en voor leerjaar 2 leerlingen uit de steekproef.

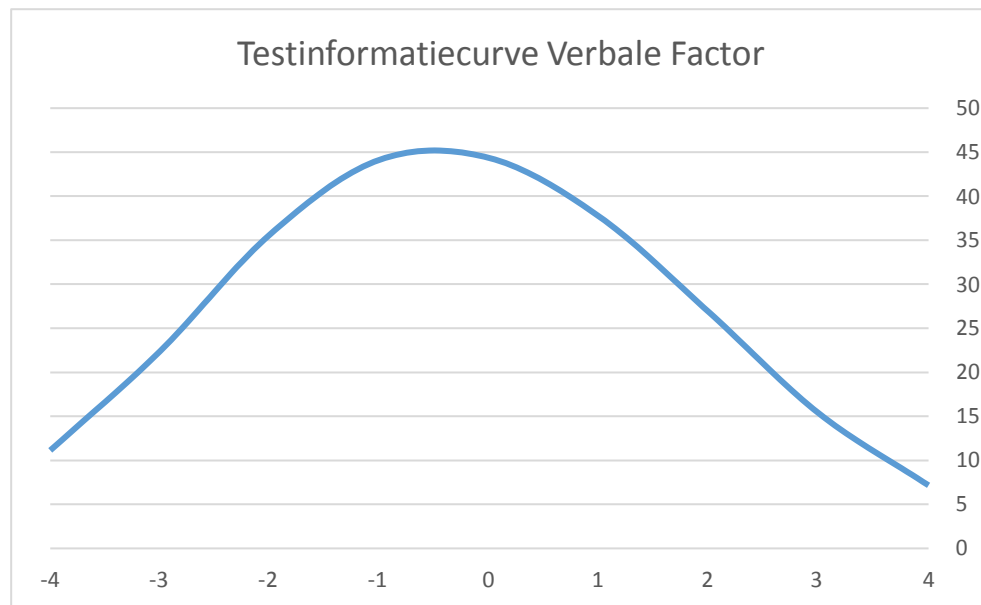


Figuur 3. Normaalverdeling TIQ-scores bij verschillende groepen.

B9. Testinformatiecurve

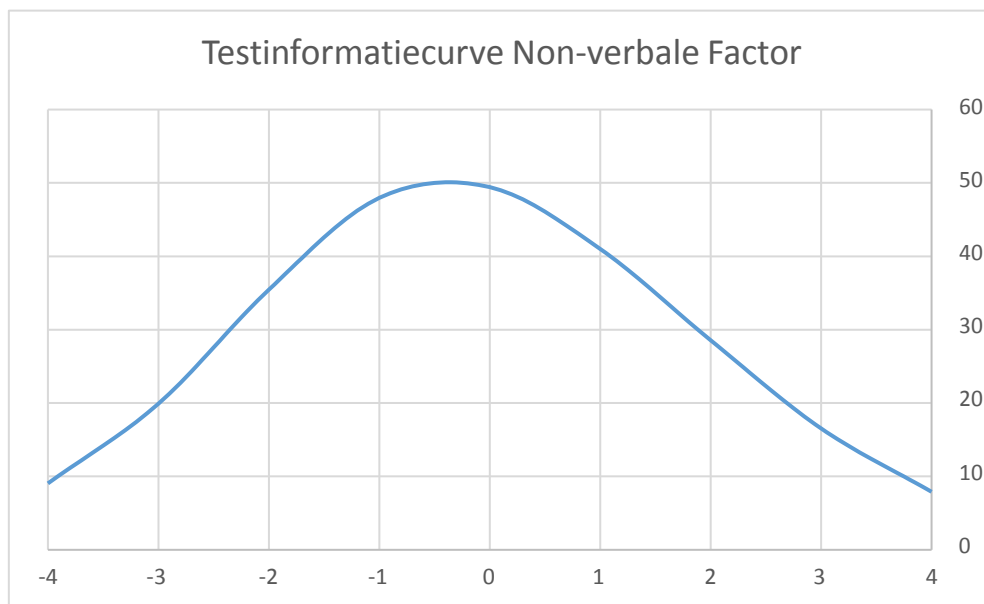
Aansluitend aan het onderzoek is er aan de hand van de items uit de huidige itembank een testinformatiecurve opgesteld. Deze is gemaakt voor het verbale, non-verbale en de totale subtesten. De testinformatiecurve geeft weer voor welk theta niveau de test het meeste informatie geeft. Dit is in de grafiek terug te vinden door af te lezen waar binnen de grafiek zich een piek bevindt. Voor deze theta-waarde geeft de test het meeste informatie.

Testinformatiecurve verbale factor



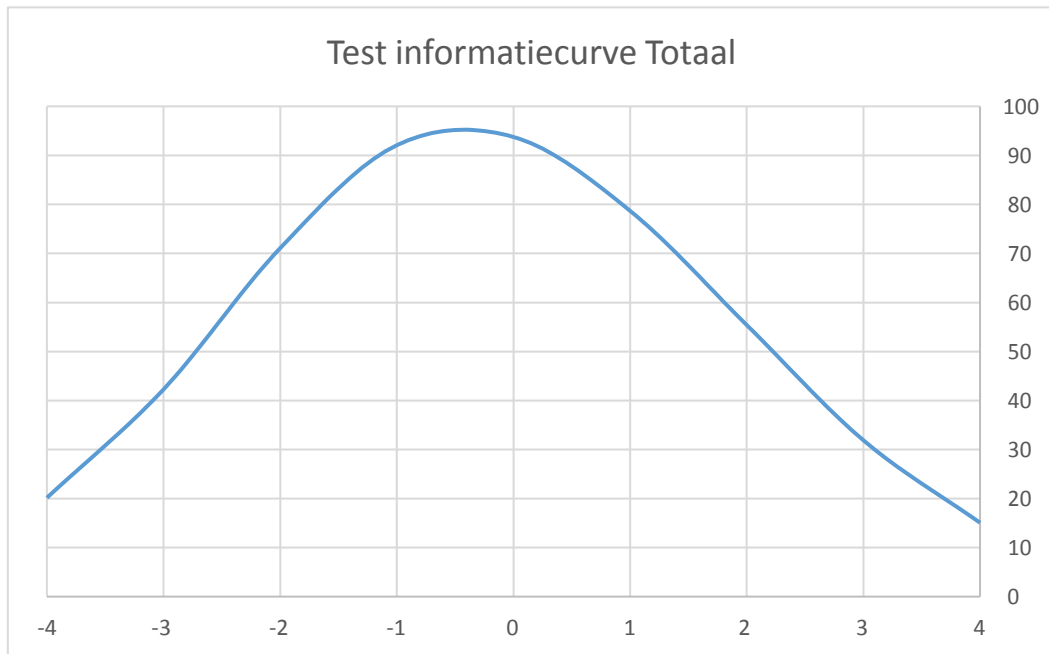
Figuur 4. Testinformatiecurve van de verbale subtesten.

Testinformatiecurve Non-verbale factor



Figuur 5. Testinformatiecurve van de non-verbale subtesten.

Test informatiecurve Totaal



Figuur 6. Testinformatiecurve van de ADIT.

B10. Uitkomsten enquête

In onderstaande tabel zijn de items van de enquête verder uitgewerkt.

Tabel 17 Frequentie, gemiddelde, modus, standaarddeviatie en de range van de items op de enquête

	Niet mee eens	Een beetje mee eens	Mee eens	Helemaal mee eens	Gemiddeld	STD	Modus	Range
De testafname van de ADIT vond ik positief	6,4%	28,2%	58,7%	6,7%	2,66	0,7	3	1-4
Het feit dat de test op een computer wordt afgenomen vind ik positief	3,9%	12,3%	39,9%	43,9%	3,24	0,814	4	1-4
Dat de test op mijn niveau wordt afgestemd vind ik positief.	5%	8,4%	44,1%	42,5%	3,24	0,808	3	1-4
De tijdsduur van de test vond ik goed.	28,2%	42,5%	22,1%	7,3%	2,08	0,888	2	1-4
Ik kon mijn concentratie aan het begin van de test er goed bijhouden	7,8%	16,5%	50,3%	25,4%	2,93	0,854	3	1-4
Aan het einde van de test kon ik mijn concentratie er ook nog goed bijhouden.	18,2%	35,8%	36,3%	9,8%	3,38	0,892	3	1-4
Ik had voldoende tijd om een vraag binnen de tijd in te vullen.	20,1%	35,5%	34,1%	10,3%	2,35	0,915	2	1-4
Ik vond de instructie voorafgaand aan de vragen helder en duidelijk.	3,1%	12,3%	41,6%	43%	3,25	0,786	4	1-4
Ik begreep wat ik moest doen.	2,2%	8,9%	46,4%	42,5%	3,29	0,721	3	1-4
De woordkeuze in de test vond ik duidelijk en goed	5,3%	18,7%	51,4%	24,6%	2,95	0,803	3	1-4
De manier waarop de vragen gesteld worden vond ik duidelijk en goed.	3,4%	12,6%	58,1%	26%	3,07	0,719	3	1-4
De plaatjes gebruikt in de test vond ik helder en duidelijk	5,6%	13,1%	52%	29,3%	3,05	0,805	3	1-4
Ik wist waar ik informatie moest vinden en waar ik moest klikken om verder te komen.	2,5%	5,3%	43,9%	48,3%	3,38	0,703	4	1-4

B11. Dankwoord scholen

Op deze plaats wil ik de scholen die met hun leerlingen hebben meegewerkt aan dit onderzoek graag bedanken. Zonder de tijd en moeite die de betreffende leerkrachten in de afname van de ADIT hebben gestoken, en zonder de inzet van hun leerlingen had dit onderzoek niet plaats kunnen vinden.

Daarom wil ik graag, als afsluiter de volgende scholen, hun medewerkers en leerlingen bedanken:

CSG Bogerman

CSG Reggesteyn

Piter Jelles de Foorakker

Het Echnaton

CSG Het Noordik: Vroomschoop

SSG De Rede

Twents Carmel College: Denekamp

Ubbo Emmius: Winschoten

