

Niet-verbale intelligentietests



Een onderzoek naar de WVN-NL en de SON-R.

Sebastiaan Abbink
Yosta Oude Weernink
Deventer, 15 augustus 2010

Niet-verbale intelligentietests

Een onderzoek naar de Wechsler Non Verbal - Nederlandse vertaling
en de Snijders Oomen niet-verbale intelligentietest - Revised

Sebastiaan Abbink
Yosta Oude Weernink
Deventer, 20 augustus 2010

IJsselgroep Educatieve Dienstverlening
Mascha Schnitzler
Bachlaan 11
7002 MZ Doetinchem

Saxion Hogeschool Deventer
Academie Mens & Arbeid

Janneke Jolij
Jaap Veldhuijzen

Voorwoord

Ter afsluiting van de opleiding tot psychodiagnostisch werker van de Academie Mens en Arbeid, onderdeel van het Saxion te Deventer is deze scriptie geschreven. De afstudeeropdracht is uitgevoerd in opdracht van de IJsselgroep te Doetinchem.

De scriptie is een weergave van een onderzoek dat gedaan is naar de WNV-NL en de SON-R. Daarnaast is het een middel om verantwoording af te leggen aan de opleiding en vormt het de basis voor het eindgesprek ter afsluiting van de opleiding. Verder kan de scriptie gebruikt worden als leidraad voor vervolgonderzoek en dient het ter naslagwerk voor studenten en geïnteresseerden.

Van de IJsselgroep willen wij graag Mascha Schnitzler bedanken voor het mogelijk maken van deze opdracht. Daarnaast willen wij Ginny Roeloffzen en Leo Schipperheijn bedanken voor hun bijdragen aan het onderzoek. Tevens gaat onze dank uit naar Vera Kneplé en Ben Wanders van het Metzo college voor de begeleiding tijdens de afnames bij de leerlingen.

Verder willen wij onze eerste begeleidster, Janneke Jolij, bedanken voor de tijd en energie die zijn in de begeleiding van onze opdracht heeft gestoken. Dank gaat ook uit naar Jaap Veldhuijzen voor zijn begeleiding en adviezen op het inhoudelijke vlak de scriptie. Tot slot bedanken wij Guido Roemer voor zijn hulp met betrekking tot de statistische analyses.

Sebastiaan Abbink
Yosta Oude Weernink

Deventer, augustus 2010

INHOUDSOPGAVE

	Pagina
SAMENVATTING	6
VERKLARENDE WOORDENLIJST	7
1. INLEIDING	9
1.1 De opdrachtgever	9
1.2 Aanleiding van het onderzoek	9
1.3 Doel van het onderzoek en probleemstelling	10
1.4 Opzet van de scriptie	11
2. THEORETISCHE ACHTERGROND	12
2.1 Intelligentie	12
2.2 Intelligentietests	13
2.3 Niet-verbale intelligentietests	15
2.3.1 Snijders-Oomen niet-verbale intelligentietests - Revised (SON-R)	15
2.3.2 Wechsler Non Verbal - NL (WNV-NL)	16
2.4 Cultuur en de invloed op intelligentie	17
2.5 Het Flynn Effect	18
2.6 ISK	19
3. METHODEN VAN ONDERZOEK	21
3.1 Onderzoeksopzet	21
3.2 Steekproef	21
3.3 Gevolgde procedure	21
3.4 Analyses	22
3.4.1 Pearson PMCC	22
3.4.2 T-toets	23
3.4.3 Standaarddeviaties en gemiddelden	23
4. RESULTATEN VAN HET ONDERZOEK	24
4.1 Steekproefkenmerken	24
4.2 Toetsing deelvragen en hoofdvraag	25
4.2.1 Uitwerking deelvraag één	25
4.2.2 Uitwerking deelvraag twee	26
4.2.3 Uitwerking deelvraag drie	26
4.2.4 Uitwerking deelvraag vier	27
5. CONCLUSIES EN AANBEVELINGEN	29
5.1 Conclusies	29
5.1.1 Conclusie deelvraag één	29
5.1.2 Conclusie deelvraag twee	29
5.1.3 Conclusie deelvraag drie	30

5.1.4	Conclusie deelvraag vier	30
5.1.5	Conclusie hoofdvraag	31
5.1.6	Overige bijzonderheden	31
5.2	Aandachtspunten betreffende het onderzoek	31
5.3	Vervolgonderzoek	32
5.4	Aanbevelingen	32
6.	LITERATUURLIJST	34
7.	BIJLAGEN	36
Bijlage 1	Beschrijving categorieën RVC	37
Bijlage 2	Interview met Leo Schipperheijn	38

SAMENVATTING

In het kader van de afstudeeropdracht van de opleiding tot psychodiagnostisch werker van het Saxion, locatie Deventer is een onderzoek verricht voor en in opdracht van de IJsselgroep te Doetinchem. Het onderzoek is uitgevoerd om te bepalen of de WNV-NL test een geschikte vervanger is van de SON-R test. Omdat de WNV-NL niet bekend is bij de IJsselgroep en de SON-R gedateerd is, wil de IJsselgroep graag weten wat de voor- en nadelen van beide tests zijn.

Om te kunnen bepalen of de WNV-NL een geschikte vervanger voor de SON-R is, zijn de tests bij dezelfde groep leerlingen afgenomen om zo de correlatie te bepalen tussen de twee tests. Hiermee kan onderzocht worden of er sprake is van soortgenootvaliditeit tussen de WNV-NL en de SON-R. Daarnaast is er bepaald of beide tests voldoen aan de voorwaarden die gesteld zijn door de minister van OC&W. Verder zijn enkele praktische aspecten getoetst, zoals de afnameduur, de normen van de test en de invloed van het Flynn effect op de normen. Tot slot is bepaald of de tests opdrachten bevatten met cultuurgevoelige en verouderde items. De steekproef bestond uit 32 leerlingen uit ISK klassen van het Metzo college. De leerlingen zijn door het Metzo college aangemeld bij de IJsselgroep om een indicatie te krijgen van hun intelligentie.

Op basis van de onderzoeksresultaten kan geconcludeerd worden dat de WNV-NL een geschikte vervanger is voor de SON-R. Dit omdat de WNV-NL voldoet aan de voorwaarden die gesteld zijn door de minister van OC&W. Daarnaast heeft de WNV-NL het voordeel dat de afnameduur korter is dan die van de SON-R. Een WNV-NL afname duurt gemiddeld 35 minuten en een afname van de SON-R twee uur. Daarnaast zijn de normen van de WNV-NL nog niet onderhevig aan het Flynn effect. De samenhang tussen de totaalscores van de WNV-NL en de SON-R is sterk ($r=.88$). Dit betekent dat er sprake is van soortgenootvaliditeit tussen beide tests en dat ze waarschijnlijk een goede indicatie geven van de niet-verbale intelligentie bij deze doelgroep. Ook is er geconstateerd dat er geen nadelig effect is door invloeden van cultuurgevoelige en verouderde items. De opdrachten die dergelijke items bevatten, correleren sterker met elkaar dan de opdrachten met neutrale, betekenisloze inhoud. De opdrachten met de cultuurgevoelige en verouderde inhoud scoren gemiddeld hoger dan de opdrachten die geen cultuurgevoelige en verouderde items bevatten. De verschillen zijn niet significant, maar de constatering dat ze gemiddeld wat hoger scoren is wel interessant, omdat blijkt dat de leerlingen beter scoren op opdrachten met concrete inhoud dan op opdrachten met abstracte inhoud.

Aan de IJsselgroep wordt aanbevolen om de SON-R vervangen wordt door de WNV-NL. De WNV-NL is een geschikte vervanger omdat de test sterk correleert met de uitslagen van de SON-R. De WNV-NL voldoet aan de voorwaarden van de minister van OC&W en heeft een kortere afnameduur dan de SON-R. De SON-R is een verouderde test, met als grootste nadelen de lange afnameduur en de verouderde normen. De normen van de SON-R zijn ouder dan 20 jaar en zullen van de COTAN de beoordeling onvoldoende meekrijgen. Hiermee voldoet de test niet meer aan de eisen van de minister en zal de test van de lijst van toegestane instrumenten worden gehaald.

VERKLARENDE WOORDENLIJST

Begripsvaliditeit	De accuraatheid waarmee een test de psychologische processen meet die binnen een theorie gespecificeerd worden.
Betrouwbaarheid	De precisie en stabiliteit van de meting, of de mate waarin het resultaat van de meting tot stand is gekomen, zonder beïnvloed te zijn door toevalsfactoren.
Bias	Vooringenomenheid
Correlatie	De mate waarin twee variabelen met elkaar samenhangen
Correlatiecoëfficiënt	Getal dat een indicatie geeft van de samenhang tussen variabelen
COTAN	Commissie Testaangelegenheden Nederland
Criteriumvaliditeit	Een maat die aangeeft in welke mate de test scores correleren met een andere maat voor de vaardigheid die men wil meten.
Cultuur faire test	Een test die geschikt is voor alle culturen waarvoor de test gemaakt is
Cultuur vrije test	Een test die geheel vrij is van cultuurafhankelijke inhoud
Empirisch	Op waarneming, ervaring en onderzoek berust.
IQ	Intelligentiequotiënt
Normen	De test scores die uit een grote en representatieve normgroep verkregen zijn, waarmee de test scores van een individu vergeleken kunnen worden
P-waarde	De overschrijdingskans van een steekproefuitkomst

Significant	Zich onderscheidend van het toeval; zodanig bijzonder onder toevalsomstandigheden, dat het betreffende verschijnsel waarschijnlijk niet alleen aan het toeval kan worden toegeschreven.
SON-R	Snijders Oomen niet verbale intelligentietest - Revised. In dit onderzoek is de versie gebruikt voor de leeftijden van 5,5 tot en met 17 jaar.
Soortgenootvaliditeit	Een hoge correlatie tussen twee tests met dezelfde meetpretentie
Steekproef	Een groep personen die representatie zijn voor de populatie
WNV-NL	Wechsler Nonverbal, Nederlandse vertaling. In dit onderzoek is de versie gebruikt voor de leeftijden van 8 tot en met 21 jaar.

1. INLEIDING

In dit hoofdstuk zullen de opdrachtgever en de aanleiding van het onderzoek beschreven worden. Verder volgt een beschrijving van de probleemstelling en het doel van het onderzoek, ten slotte wordt de opzet van de scriptie behandeld.

1.1 Opdrachtgever

De IJsselgroep is een organisatie die zich richt op het onderwijs. De opdrachtgever van dit onderzoek is de afdeling Educatieve Diensverlening. Deze afdeling richt zich op psychologische dienstverlening voor basisscholen en middelbare scholen die zijn aangesloten bij de IJsselgroep. Wanneer er bij de leerlingen van deze scholen leer- of gedragsproblemen ontstaan, kan de IJsselgroep door middel van een psychologisch onderzoek duidelijkheid geven over de aard en de ernst van de problemen. Daarbij spelen niet alleen de leerproblemen van het kind een rol, maar ook de ontwikkelingsmogelijkheden. Bij de afdeling Educatieve Dienstverlening werken g-z-psychologen, orthopedagogen en psychodiagnostisch medewerkers. Kinderen in de leeftijd van 4 tot 18 jaar kunnen aangemeld worden voor een psychologisch onderzoek. De basis- en middelbare scholen die aangesloten zijn bij de IJsselgroep melden het merendeel van de onderzoeken aan, maar particulieren kunnen ook een onderzoek aanvragen. Naast het psychologisch onderzoek biedt de IJsselgroep ook verschillende soorten van begeleiding aan, zoals dyslexiebehandelingen en gedragsbegeleiding.

1.2 Aanleiding van het onderzoek

Eens per jaar wordt er door de IJsselgroep een grote groep leerlingen onderzocht die uit internationale schakelklassen (ISK's) komen. De aanmelding voor deze onderzoeken komt van het Metzo college te Doetinchem, dat ook een dependance in Winterswijk heeft. Het Metzo College is een school voor het voorbereidend middelbaar beroepsonderwijs (VMBO) en voor het beroepsgerichte middelbaar algemeen voortgezet onderwijs (MAVO). De leerlingen die in deze ISK klassen zitten, zijn van buitenlandse afkomst en wonen vaak nog maar een jaar in Nederland. Zodra deze kinderen naar Nederland komen, zijn zij leerplichtig (Rijksoverheid, 2010a). Het reguliere onderwijs is echter niet geschikt voor hen, omdat ze niet of nauwelijks de Nederlandse taal beheersen. Voordat ze naar het reguliere onderwijs kunnen, zullen ze eerst de Nederlandse taal voldoende moeten beheersen. Het volgen van onderwijs zal anders heel moeilijk zijn voor deze leerlingen als ze de docenten niet begrijpen en schoolboeken niet kunnen lezen. Daarnaast is het bij deze kinderen onduidelijk in hoeverre ze in hun thuisland onderwijs hebben gevolgd en wat hun capaciteiten zijn, daarom worden ze eerst opgevangen in een ISK klas (Schipperheijn, persoonlijke communicatie, 7 juni 2010). Vanuit de ISK klas wordt bepaald welke vorm van onderwijs het beste bij deze leerlingen past. Het is belangrijk om de leerlingen op de goede plek te krijgen zodat problemen voorkomen kunnen worden, zoals uitval van leerlingen die op een te hoog niveau zitten en onderpresteren van leerlingen die op een te laag niveau zitten (Schipperheijn, persoonlijke communicatie, 7 juni 2010). Daarom wordt er bepaald of de leerlingen tot hun recht komen op het leerwegondersteunend onderwijs (LWOO) of het praktijkonderwijs (PrO). Het leerwegondersteunend onderwijs is een vorm van ondersteuning die een leerling kan krijgen wanneer hij een van de vier opleidingen volgt van het VMBO (Rijksoverheid,

2010b). Leerlingen waarvoor het VMBO en het LWOO niet geschikt is, komen in aanmerking voor het praktijkonderwijs. Deze opleiding leidt leerlingen direct op voor de arbeidsmarkt (Rijksoverheid, 2010c).

1.3 Doel van het onderzoek en probleemstelling

Tijdens de onderzoeken van de ISK leerlingen, of ISK onderzoeken, wordt er bij de leerlingen een intelligentietest afgenomen. Op dit moment gebruikt de IJsselgroep de Snijders-Oomen niet-verbale intelligentietest 5,5-17 (SON-R) voor de ISK onderzoeken (Schipperheijn, persoonlijke communicatie, 7 juni 2010). De vraag van de IJsselgroep is of de Wechsler Non Verbal-NL (WNV-NL) de SON-R zou kunnen vervangen (Schnitzler, persoonlijke communicatie, 11 februari 2010). Het is belangrijk voor de IJsselgroep dat er veel afnames gedaan kunnen worden in korte tijd. De afnameduur van de WNV-NL is korter dan die van de SON-R en daarom wil de IJsselgroep weten of de WNV-NL inzetbaar is bij de ISK onderzoeken. Omdat de WNV-NL in 2008 op de markt is gekomen, is het nog een vrij nieuwe test. De IJsselgroep wil graag weten wat de ervaringen zijn met de WNV-NL. De hoofdvraag van dit onderzoek luidt: *'Is de WNV-NL een geschikte vervanger van de SON-R?'* Om deze vraag te beantwoorden zullen een aantal deelvragen getoetst worden.

De eerste deelvraag luidt: *'Voldoen de WNV-NL en de SON-R aan de voorwaarden die gesteld zijn door de minister van Onderwijs, Cultuur en Wetenschap (OC&W)?'* Het ministerie van OC&W heeft een aantal voorwaarden gesteld aan de onderzoeksinstrumenten die gebruikt mogen worden bij de ISK onderzoeken, namelijk de test moet geschikt zijn voor de doelgroep, de test goedgekeurd moet zijn door de COTAN en de test goed bevonden moet zijn door de regionale verwijzingscommissie (RVC) (Cals & Marijnen, 1963). Het is van belang dat de tests hieraan voldoen, voordat een test mag worden ingezet bij de ISK onderzoeken.

De tweede deelvraag is: *'Wat zijn de voor- en nadelen van de praktische aspecten van de WNV-NL en de SON-R?'* Om deze vraag te beantwoorden zullen de praktische aspecten van beide tests tegenover elkaar worden gezet om te bepalen welke test handiger is in gebruik. Zo zullen de afnameduur en het bereik van de normen aan bod komen. Bij de normen is ook het Flynn effect van belang. Het Flynn effect stelt dat de normen van tests verouderen per generatie en dat deze normen zonder correctie niet meer bruikbaar zijn (Deary, 2001). Er zal bepaald worden in welke mate het Flynn effect een rol speelt bij de normen van de SON-R en de WNV-NL.

Als derde deelvraag komt aan bod: *'In hoeverre hangen de uitslagen van de WNV-NL en de SON-R samen?'* De SON-R en de WNV-NL hebben allebei dezelfde meetpretentie, dit houdt in dat ze beide de niet-verbale intelligentie van een persoon pretenderen te meten. De verwachting is dat de IQ scores van de SON-R en de WNV-NL hoog met elkaar zullen correleren en dat het correlatiequotiënt minimaal .75 zal zijn (Baarda, De Goede & Van Dijkum, 2007). Wanneer de correlatie hoger is dan .75 is er sprake van soortgenootvaliditeit.

De vierde deelvraag luidt: *'In hoeverre spelen cultuurgevoeligheid en veroudering van de test items een rol in de uitslagen?'* Er zal onderzocht worden in hoeverre de opgaven van de SON-R en de WNV-NL een invloed hebben op de scores. Het is moeilijk om de cultuurgevoeligheid en veroudering van de items apart te onderzoeken. Echter,

volgens Tellegen & Laros (2004) is het ook niet nodig om ze apart te onderzoeken omdat zowel de cultuurgevoeligheid als de veroudering een nadelig effect kan hebben op de scores. In dit onderzoek zullen de cultuurgevoeligheid en veroudering van de items samen onderzocht worden. Om de deelvraag te beantwoorden zullen de correlaties bepaald worden tussen de subtests *stripverhalen - plaatjes ordenen* en *analogieën - matrix redeneren*. De onderdelen *stripverhalen* van de SON-R en *plaatjes ordenen* van de WNV-NL zijn vergelijkbare opdrachten. Bij beide opdrachten moeten plaatjes in de goede volgorde gelegd worden om zo een logisch verhaal te vormen. Ze doen allebei een beroep op inzicht in sociale situaties. Beide subtests bevatten items met concrete afbeeldingen. De onderdelen *analogieën* (SON-R) en *matrix redeneren* (WNV-NL) doen beiden een beroep op het vermogen om abstract te redeneren. De items in deze opdrachten bevatten alleen geometrische figuren (abstracte afbeeldingen) en zijn dus minder gevoelig voor culturele bias (Tellegen & Laros, 2004). De correlatiequotiënten kunnen worden vergeleken en daarna kan worden bepaald in hoeverre ze corresponderen met de correlatie tussen de totaalscores. Wanneer er sprake is van cultuurgevoeligheid en veroudering van de items, dan is de verwachting dat de subtests met concrete inhoud minder sterk zullen correleren dan de subtests met abstracte inhoud. Daarnaast kunnen de gemiddelde scores van de subtests bepaald worden om te bepalen op welke subtests de leerlingen beter scoren. Naar verwachting zullen de leerlingen gemiddeld beter presteren op de subtests met abstracte inhoud, dan op de subtests met concrete inhoud.

1.4 Opzet van de scriptie

De scriptie bestaat uit een aantal hoofdstukken. In hoofdstuk 1 wordt het onderzoek en de vraagstelling ingeleid. Daarop aansluitend wordt in hoofdstuk 2 een theoretisch kader rondom intelligentie en intelligentietests uiteengezet. De methodiek van het onderzoek wordt in hoofdstuk 3 nader toegelicht. In hoofdstuk 4 worden de resultaten van het onderzoek beschreven met behulp van de gestelde deelvragen. In hoofdstuk 5 wordt de hoofdvraag beantwoord en zullen er kanttekeningen geplaatst worden bij het onderzoek. Ook wordt in dit hoofdstuk de mogelijkheid voor vervolgonderzoek besproken worden en zullen er aanbevelingen gedaan worden ten behoeve van de IJsselgroep.

2. THEORETISCHE ACHTERGROND

In dit hoofdstuk wordt het begrip intelligentie uitgelegd en zal er dieper op het onderwerp 'niet-verbale intelligentie' worden ingegaan. Vervolgens wordt beschreven wat intelligentietests zijn en hoe deze in het werkveld ingezet worden. Ook zal het 'Flynn effect' worden beschreven en hoe culturele achtergrond van invloed kan zijn op de intelligentietests. Tot slot wordt er aandacht besteed aan het ontstaan van internationale schakelklassen. Daarin wordt ook besproken wat ISK's precies zijn, waarom psychologische onderzoeken belangrijk zijn en welke zaken hierover in de wet beschreven zijn.

2.1 Intelligentie

Er is geen eenduidige omschrijving voor het begrip 'intelligentie'. Volgens Brysbaert (2006) is intelligentie op te delen in drie soorten, namelijk de analytische intelligentie, de praktische intelligentie en de sociale en emotionele intelligentie. In dit onderzoek is de analytische intelligentie van belang en deze kan opgevat worden als: het geheel van verstandelijke vermogens, processen en vaardigheden, die ervoor zorgen dat men abstract, logisch en consistent kan redeneren. Men is ook in staat om relaties te ontdekken, problemen op te lossen en regels te ontdekken in ogenschijnlijk ongeordend materiaal (Resing & Drenth, 2007).

Resing en Drenth (2007) veronderstellen dat er verschillende niveaus van intelligentie zijn. Zij gaan uit van de niveaus zoals Philip Vernon ze in 1969 heeft beschreven, namelijk intelligentie A, intelligentie B en intelligentie C. Volgens Vernon (in Resing & Drenth, 2007) is intelligentie A het aangeboren, genetische bepaalde potentieel tot intelligent handelen. Het is dus niet te beïnvloeden door de omgeving, door voeding of door opvoeding en leerervaringen. Intelligentie A is een theoretische veronderstelling van psychologen en is niet direct te meten of te observeren. Intelligentie B is echter wel te meten. Dit niveau van intelligentie ontstaat door de wisselwerking tussen de genetische aanleg (intelligentie A) en de omgevingsinvloeden en leerervaringen. Intelligentie B kan uitgelegd worden als de verstandelijke vermogens waarover iemand beschikt op een bepaald moment en die het resultaat zijn van alle voorafgaande (leer)ervaringen. Het laatste niveau is intelligentie C. Intelligentie C probeert een zo goed mogelijke schatting te geven van intelligentie B met behulp van een intelligentietest. Intelligentie C is daarmee een maat voor de intelligentie die uitgedrukt wordt in een intelligentiequotiënt. Intelligentie B is de daadwerkelijke analytische intelligentie die een persoon heeft en intelligentie C is de gemeten intelligentie van die persoon.

In de psychologie wordt er vanuit gegaan dat de analytische intelligentie een algemene factor (g-factor) bevat. Deze g-factor speelt een rol bij het oplossen van alle intelligentietaken (Brysbaert, 2006). Dit houdt in dat wanneer een persoon goed presteert op één taak, zeer waarschijnlijk ook goed zal presteren op een andere taak (Roediger et al., 2004). Charles Spearman, die deze algemene factor ontdekte, vond ook duidelijke relaties tussen bepaalde verschillende taken. Hij stelde, dat er naast de algemene factor ook specifieke mentale vaardigheden (s-factoren) aanwezig zijn, die van invloed zijn bij het oplossen van specifieke taken (Roediger et al, 2004; Brysbaert, 2006). Cattall en Horn stelden voor dat de analytische intelligentie niet bestaat uit g- en s-factoren, maar uit een vloeiende intelligentie en een gekristalliseerde

intelligentie (Cattell & Horn in Brysbaert, 2006). Echter, de samenhang tussen vloeiende en gekristalliseerde intelligentie is groot en dat zou erop kunnen wijzen dat er toch een g-factor aanwezig is (Gleitman, Fridlund & Reisberg, 1999). De hiërarchische theorie biedt op dit moment de beste verklaring voor de opbouw van de intelligentie. De hiërarchische theorie stelt dat de analytische intelligentie bestaat uit zowel een g-factor als een vloeiende en gekristalliseerde intelligentie en enkele s-factoren (Brysbaert, 2006). De specifieke factoren worden ook wel de primaire factoren genoemd en omvatten vaardigheden als verbaal vermogen, numeriek vermogen, ruimtelijk inzicht en snelheid van waarneming (Resing & Drenth, 2007). De primaire factoren kunnen in twee groepen verdeeld worden, namelijk de vloeiende en de gekristalliseerde intelligentie. Vloeiende intelligentie is een maat voor het vermogen om abstract te redeneren en de snelheid en flexibiliteit in het denken. Gekristalliseerde intelligentie meet de kennis en vaardigheden die geleerd zijn gedurende de levensloop (Gleitman, Fridlund & Reisberg, 1999; Resing & Drenth, 2007). De vloeiende intelligentie wordt gezien als veel minder gevoelig voor leerervaringen en invloed van cultuur- en omgevingsfactoren dan de gekristalliseerde intelligentie. De vloeiende intelligentie komt dus grotendeels overeen met intelligentie A en de gekristalliseerde intelligentie komt in de buurt van intelligentie B (Resing & Drenth, 2007). De vloeiende en de gekristalliseerde intelligentie samen vormen de g-factor (Brysbaert, 2006).

De leerlingen uit de ISK klassen zijn allemaal van buitenlandse afkomst en komen uit verschillende culturen. Op het moment dat het ISK onderzoek plaats vindt, hebben zij een jaar onderwijs gevolgd in Nederland. Hun kennis van de Nederlandse taal is dan nog onvoldoende. Een intelligentietest die zowel de vloeiende als de gekristalliseerde intelligentie meet zou niet op zijn plaats zijn, omdat de invloeden van cultuur en taal een negatief effect kunnen hebben op de uitslag bij de ISK leerlingen (Gleitman, Fridlund & Reisberg, 1997). Om de ISK leerlingen op een eerlijke en betrouwbare manier te testen, worden er tests ingezet die een beroep doen op de vloeiende intelligentie.

2.2 Intelligentietests

Intelligentietests zijn instrumenten die professionals in staat stellen om een indicatie te krijgen van het denkvermogen van een persoon. Bij het ontwerpen van intelligentietests wordt de hiërarchische theorie veel gebruikt, omdat deze theorie tot nu toe het best onderzocht en empirisch onderbouwd is (Resing & Drenth, 2007). De meeste intelligentietests zullen dus een beroep doen op de vloeiende en de gekristalliseerde intelligentie om zo de g-factor te bepalen (Roediger et al., 2004; Brysbaert, 2006).

Volgens Resing en Drenth (2007) is een intelligentietest een systematische meetprocedure, bestaande uit een aantal opdrachten die zorgvuldig zijn uitgezocht. De opdrachten worden op een gestandaardiseerde manier voorgelegd aan een proefpersoon en de oplossingen worden daarna op een objectieve manier verwerkt door een beoordelaar. Hieruit zijn al enkele kenmerken af te leiden die een intelligentietest een goede test maken, namelijk standaardisatie en objectiviteit. Daarbij horen ook betrouwbaarheid, validiteit en normering (Resing & Drenth, 2007). Standaardisatie van een test betekent dat de uitvoering van de testmaterialen, de instructies en de wijze waarop getest moet worden zo goed mogelijk voorgeschreven is

(Resing & Drenth, 2007). Elke afname van de test is elke keer zo veel mogelijk hetzelfde. De objectiviteit van een test wil zeggen, dat de verwerking en beoordeling van de antwoorden zo veel mogelijk zonder subjectieve invloeden van de testleider plaatsvindt (Resing & Drenth, 2007). Bij intelligentietests zorgt een gestandaardiseerde handleiding daarvoor. De handleiding bevat uitvoerige, stapsgewijze procedures met betrekking tot wat de testleider moet doen en hoe er gescoord moet worden (Resing & Drenth, 2007). Een goede test is ook betrouwbaar, hetgeen verwijst naar de stabiliteit van de testscores bij verschillende afnames (Roediger et al., 2004). Wanneer een proefpersoon bij een afname goed op een test scoort, moet hij bij een tweede afname van dezelfde test ook goed scoren (Brybaert, 2006). De validiteit van een test is ook belangrijk, want een valide test meet datgene wat hij beweert te meten (Roediger et al., 2004). Een intelligentietest die een IQ score geeft, maar geen goede indicatie geeft van de intelligentie, is niet valide. Tot slot is er de normering. De scores van mensen bij wie een intelligentietest is afgenomen, zullen vergeleken worden met de scores van een normgroep. De normgroep bestaat uit mensen van dezelfde leeftijd als de proefpersoon die onderzocht is (Roediger, et al., 2004). De normgroepen worden samengesteld door een steekproef te trekken uit een bepaalde populatie, bijvoorbeeld Nederland (Brybaert, 2006). De kwaliteit van de normgroepen is afhankelijk van de representativiteit van de steekproef (Brybaert, 2006).

Intelligentietests mogen in Nederland niet zomaar gebruikt worden. Ze moeten eerst gekeurd en goed bevonden worden door de COTAN (Brybaert, 2006). De COTAN toetst een test op zeven verschillende punten: de uitgangspunten bij de testconstructie, de kwaliteit van het testmateriaal, de kwaliteit van de handleiding, de normen, de betrouwbaarheid, de begripsvaliditeit en de criteriumvaliditeit (Evers, Lucassen, Meijer & Sijtsma, 2010). De beoordeling van de uitgangspunten bij de testconstructie stelt vast of het gebruiksdoel, de theoretische achtergrond en de operationalisatie daarvan zijn beschreven in de testinhoud (Evers et al., 2010). Bij de kwaliteit van het testmateriaal beoordeelt de COTAN de standaardisatie van de testopgaven, de scoring en de instructies (Evers et al., 2010). Tevens wordt bepaald of items kwetsend kunnen zijn voor specifieke bevolkingsgroepen. Bij de kwaliteit van de handleiding wordt er vooral gelet op de informatie die wordt geboden voor de ondersteuning van de testleider bij de afname en de interpretatie van de test (Evers et al., 2010). De beoordeling van de normen bepaald vooral wat de kwaliteit van de normen is. Er wordt nu ook meer rekening gehouden met de verjaringstermijn van de normen. Wanneer er na vijftien jaar geen hernormeringsonderzoek heeft plaatsgevonden, krijgen de normen de kanttekening 'de normen zijn verouderd'. Na twintig jaar luidt de kanttekening 'wegens veroudering zijn de normen niet meer bruikbaar' en worden de normen als onvoldoende beoordeeld (Evers et al., 2010). De betrouwbaarheid van een test wordt uitgedrukt in een correlatiecoëfficiënt. Van tests die voor belangrijke beslissingen worden ingezet, wordt verwacht dat de test een correlatie van minimaal .9 heeft (Evers et al., 2010). Volgens Evers et al. (2010) zijn belangrijke beslissingen bijvoorbeeld personeelsselectie, verwijzing naar speciaal onderwijs en opname of ontslag in een kliniek. Bij ISK onderzoeken worden intelligentietests ingezet om een verwijzing te krijgen voor het speciaal onderwijs, hiermee horen deze onderzoeken bij de belangrijke beslissingen. Tot slot geeft de COTAN een beoordeling over de begripsvaliditeit en de criteriumvaliditeit. Begripsvaliditeit is de precisie waarmee een test de psychologische processen meet die binnen de gestelde theorie genoemd worden (Brybaert, 2006). De criteriumvaliditeit meet de correlatie tussen de

testscores en een andere maat voor dezelfde vaardigheid (Brybaert, 2006). Bij intelligentietests kunnen IQ scores vergeleken worden met schoolresultaten om te bepalen of een hoger IQ ook leidt tot hogere schoolprestaties.

2.3 Niet-verbale intelligentietests

Intelligentietests bestaan uit opdrachten die een beroep doen op specifieke vaardigheden. De vaardigheden vormen de vloeiende en gekristalliseerde intelligentie en de algemene intelligentie (Brybaert, 2006). Bij niet-verbale intelligentietests proberen testmakers de invloeden van taal zo veel mogelijk te verminderen. Dit kan door de verbale opdrachten uit een test te halen en door het mogelijk te maken dat instructies niet verbaal gegeven hoeven te worden (Resing & Drenth, 2007). Door middel van handgebaren en ja-knikken en nee-schudden kunnen opdrachten uitgelegd worden aan een proefpersoon, zonder daarbij woorden te gebruiken. De precieze manier waarop deze niet verbale instructies gegeven worden, staat vastgelegd in de handleiding van een test. Bij de ISK onderzoeken worden niet-verbale intelligentietests ingezet en twee van zulke tests zijn de SON-R en de WNV-NL.

2.3.1 Snijders-Oomen Niet-verbale intelligentietest - Revised (SON-R)

De SON-R is een niet-verbale intelligentietest voor kinderen van 2,5 tot 17 jaar. Er zijn twee verschillende uitvoeringen van deze test, namelijk één voor kinderen in de leeftijd van 2,5 tot en met 7 jaar en één voor kinderen van 5,5 tot en met 17 jaar (Snijders, Tellegen en Laros, 1997). In dit onderzoek is de test gebruikt voor de leeftijden van 5,5 tot en met 17 jaar. De SON-R kenmerkt zich door het geven van feedback na elke opgave. De proefpersoon krijgt na elke opgave van de testleider te horen wanneer de opgave goed of fout is. Dit stelt de proefpersoon in staat om van de fouten te leren. Hierdoor wordt ook een beroep gedaan op het vermogen om te leren, niet alleen op de algemene intelligentie op het moment van afname (Tellegen & Laros, 2004). Volgens Tellegen en Laros (2004) dragen de niet-verbale opdrachten en de feedback bij aan een cultureel faire test, want het leervermogen is onafhankelijk van de cultuur.

De SON-R bestaat uit zeven verschillende onderdelen. Deze onderdelen, ook wel subtests genoemd, zijn onder te verdelen in vier typen subtests. Ten eerste zijn er de abstracte redeneertests. Dit zijn *categorieën* en *analogieën*. Deze onderdelen doen een beroep op het vermogen om relaties tussen begrippen te ontdekken die niet aan ruimte en tijd gebonden zijn. Ten tweede komen de concrete redeneertests; *situaties* en *stripverhalen*. Hierbij gaat het juist wel om mensen of objecten die in ruimte en tijd met elkaar in verband staan. De onderdelen *mozaïeken* en *patronen* behoren tot het type ruimtelijke tests. Dit type doet, net als de concrete redeneertests, een beroep op het vermogen om een relatie binnen een ruimtelijk geheel te zien. Het verschil is echter dat het bij ruimtelijke tests gaat om vormrelaties tussen stukken van een figuur. Bij concrete redeneertests gaat het juist om betekenisrelaties tussen delen van een afbeelding. Als laatste is er nog een perceptuele test, namelijk *zoekplaten*. De proefpersoon moet bij deze opdracht een figuur opzoeken dat een aantal keer verstopt zit in een zoekplaat.

De subtests worden volgens een vaste opzet afgenomen. Er wordt begonnen met het onderdeel *categorieën*. Hierbij krijgt de proefpersoon een drietal plaatjes te zien die

bij elkaar horen. De proefpersoon moet dan uit vijf andere tekeningen er twee kiezen die bij de eerste drie horen. Vervolgens komt het onderdeel *mozaïeken*. De proefpersoon krijgt een kader waarin rood/witte blokjes gelegd kunnen worden. Met de blokjes kunnen een aantal patronen nagemaakt worden. Daarna volgt de subtest *zoekplaten* waarbij de proefpersoon een plaat te zien krijgt waarin vijftien keer een afbeelding verstopt zit. Na deze opdracht komt het onderdeel *patronen*. De proefpersoon moet een onderbroken lijnenpatroon invullen zodat de lijnen op dezelfde wijze doorlopen. Hierna volgt *situaties*. De proefpersoon kijkt naar een afbeelding van een situatie waaruit een of meerdere delen ontbreken. Uit een aantal mogelijkheden worden dan de goede oplossingen gekozen om de situatie weer tot een samenhangend geheel te maken. *Analogieën* is dan het volgende onderdeel en daarbij krijgt de proefpersoon een geometrisch figuurtje te zien dat een verandering ondergaat. De verandering moet op dezelfde manier op een ander figuur worden toegepast om het juiste antwoord te verkrijgen. De test wordt afgesloten met *stripverhalen*. De proefpersoon krijgt een aantal kaartjes waar afbeeldingen op staan. De afbeeldingen vormen een verhaal, maar liggen nog niet in de goede volgorde. Het is aan de proefpersoon om de kaartjes zo te leggen dat er een logisch verhaal gevormd wordt. De afnametijd van de test bedraagt gemiddeld anderhalf uur. Dit kan echter uitlopen tot twee uur omdat de onderdelen *categorieën*, *situaties* en *analogieën* geen maximum tijd hebben. Er wordt dus niet afgebroken wanneer een bepaalde tijd is verstreken. Nadat alle subtests zijn afgenomen, kunnen de normscores berekend worden. Aan de hand van de normscores wordt het intelligentiequotiënt bepaald voor de niet verbale intelligentie en voor de vier typen subtests.

De SON-R is in 1989 op de markt gebracht en is in 1990 door de COTAN gekeurd. De SON-R is op alle punten goed bevonden (Roeloffzen, persoonlijke communicatie, 1 april 2010). Echter, als er in twintig jaar niks verandert aan een test, wordt hij uit het bestand van de COTAN gehaald (Schipperheijn, persoonlijke communicatie, 7 juni 2010). Wanneer er een nieuw normeringsonderzoek plaatsvindt, zal een test opnieuw onderzocht worden door de COTAN en opgenomen worden in hun bestand (Evers, Lucassen, Meijer & Sytsma, 2010). Bij de SON-R is geen hernormeringsonderzoek uitgevoerd. Daarnaast zit de test twintig jaar in het bestand van de COTAN.

2.3.2 Wechsler Non Verbal - NL (WNV-NL)

De WNV-NL is de Nederlandse vertaling van de Wechsler Nonverbal Scale of Ability. De WNV-NL is een nieuwe test die onlangs op de markt is gekomen, namelijk in 2008. De WNV-NL meet de niet-verbale intelligentie van kinderen en adolescenten in de leeftijd van 4 tot en met 21 jaar. Het is met deze test mogelijk om de niet-verbale intelligentie te bepalen bij kinderen en adolescenten met communicatieve handicaps of een onvoldoende beheersing van de Nederlandse taal (WNV-NL, n.d.). Bijvoorbeeld bij kinderen met een allochtone achtergrond, autistische kinderen, maar ook bij doven en slechthorenden en kinderen met spraak- en taalproblemen. Daarnaast is de test ook algemeen bruikbaar bij kinderen die in Nederland zijn opgegroeid en de Nederlandse nationaliteit hebben. De test wordt individueel afgenomen en heeft een afnametijd van gemiddeld drie kwartier. De WNV-NL is in 2009 gekeurd door de COTAN. De test is goed bevonden op alle punten met uitzondering van de betrouwbaarheid en de criteriumvaliditeit. Deze punten hebben een voldoende gekregen. Bij de betrouwbaarheid heeft de COTAN wel de opmerking geplaatst dat de betrouwbaarheid

van de hele test voldoende is, maar de betrouwbaarheid van de subtests is onvoldoende (Roeloffzen, persoonlijke communicatie, 1 april 2010).

Hieronder worden alleen de subtests vermeld die gebruikt zijn voor dit onderzoek. De vier subtests die in dit onderzoek gebruikt zijn, zijn voor de leeftijd van 8 tot en met 21 jaar. Voor de leeftijd van 4 tot en met 8 jaar worden twee subtests toegevoegd en twee andere weggelaten (Wechsler & Naglieri, 2008). De eerste subtest is *matrix redeneren*. Het doel van deze subtest is dat de proefpersoon gaat ontdekken hoe verschillende vormen en geometrische elementen, ruimtelijk of logisch met elkaar samenhangen. De opdracht bestaan uit verschillende soorten items, bijvoorbeeld geometrische patronen, redeneren op basis van analogie en ruimtelijke visualisaties. De proefpersoon moet bij elk item het beste antwoord kiezen uit de verschillende mogelijkheden om de opdracht compleet te maken. Op basis van deze subtest kan het perceptueel redeneervermogen en de simultane informatieverwerking van de proefpersoon bepaald worden. De tweede subtest die wordt afgenomen is *substitutie B*. Bij deze subtest moet de proefpersoon in twee minuten symbolen natekenen die zijn gekoppeld aan cijfers. De proefpersoon krijgt een opgavenblad met daarop de cijfers en de bijbehorende figuren. Daaronder staan alleen de cijfers en de proefpersoon moet de figuren bij de juiste cijfers neerzetten. Deze subtest meet de verwerkingsnelheid en de grafomotorische snelheid. Daarna wordt de subtest *ruimtelijke oriëntatie* afgenomen. In deze subtest wijst de testleider een aantal blokjes aan die de proefpersoon in dezelfde en vervolgens in omgekeerde volgorde moet herhalen. Deze subtest meet het werkgeheugen voor visueel ruimtelijke stimuli. De vierde en laatste subtest die wordt afgenomen is *plaatjes ordenen*. Bij deze subtest krijgt de proefpersoon een aantal kaartjes met een afbeelding erop voor zich gelegd. Deze afbeeldingen moeten een verhaal gaan vormen die de proefpersoon in de juiste volgorde moet gaan leggen. De proefpersoon kan bij deze test bonuspunten krijgen door de opgave binnen een bepaalde tijd te maken. Deze subtest doet een beroep op de perceptuele organisatie (Wechsler & Naglieri, 2008).

2.4 Cultuur en de invloed op intelligentie

De ISK leerlingen zijn afkomstig uit verschillende landen. Dat betekent dat ze uit verschillende culturen komen en die cultuur is van invloed op de ontwikkeling van een leerling. Volgens Hofstede (1997) is cultuur te omvatten in twee verschillende definities. Doorgaans wordt met cultuur de term ‘beschaving’ bedoeld en ‘verfijning van de geest’. Vaak wordt er ook gedoeld op de producten van deze verfijning zoals educatie, kunst en literatuur. Cultuur kan echter ook opgevat worden als de patronen van het denken, voelen en doen van personen. Deze vorm van cultuur is gedeeld door een grote groep mensen die met elkaar in dezelfde sociale omgeving wonen of hebben gewoond (Hofstede, 1997). Cultuur is aangeleerd en niet genetisch bepaald. Dit leerproces begint binnen de familie en vervolgt zich in de wijk waar iemand woont, op school en op het werk (Hofstede, 1997). De cultuur die heerst binnen de familie, wijk, school en werk, beïnvloedt de nadruk die gelegd wordt op opvoeding, leefomstandigheden, onderwijs en cognitieve vaardigheden (Resing & Drenth, 2007). In culturen waar onderwijs en cognitieve vaardigheden minder belangrijk worden gevonden, ligt er een minder grote nadruk op scholing (Driscoll, 2005). Zo is onderzocht dat ongeschoolde kinderen geen onderscheid kunnen maken tussen hun gedachten en waar ze op dat moment aan dachten. Ze konden de vraag ‘Waarom denk je dat X het antwoord is?’ niet beantwoorden, maar wel wanneer hij geformuleerd

werd als ‘Waarom is X het antwoord?’ (Greenfield, in Driscoll, 2005). Bij het gebruik van een intelligentietest met verbale opdrachten, zouden de vragen op een manier geformuleerd zijn die sommige leerlingen niet begrijpen. Het zou dus kunnen zijn dat een leerling het antwoord op een vraag wel weet, maar vanwege de vraagstelling niet op het antwoord kan komen. Dit zou de leerling punten kunnen kosten. In westerse culturen ligt de nadruk op tijd en snelheid. Andere culturen delen deze nadruk niet en leggen de focus meer op goed overwogen beslissingen maken. Wanneer ze een antwoord op een vraag niet weten, zullen ze vaker geen antwoord geven dan gokken naar het antwoord (Gleitman, Fridlund & Reisberg, 1999). Dit kan een nadelige invloed hebben op de prestaties op intelligentietests, omdat snelheid en gokken naar het antwoord ook punten op kunnen leveren. Bij de subtest *plaatjes ordenen* van de WNV-NL moet de leerling binnen een bepaalde tijd de opdracht goed maken. Het is dus van belang dat de leerling vlot aan de opdracht kan werken. Daarnaast kan de leerling bonuspunten krijgen door de opgaven sneller af te hebben dan de maximum voorgeschreven tijd. Bij de subtest *substitutie B* is de taak dat de leerling zo veel mogelijk items goed maakt binnen een bepaalde tijd. Ook hier ligt de nadruk dus op snelheid. Bij de SON-R zijn de maximum tijden wat ruimer, de leerling mag wat langer over een opgave doen dan bij de subtests van de WNV-NL. De subtest *stripverhalen* van de SON-R heeft voor alle opgaven een maximum tijd van drie minuten. Dezelfde subtest van de WNV-NL, *plaatjes ordenen*, heeft een maximum tijd van 45 seconden en naarmate de opgaven moeilijker worden een maximum tijd van 90 seconden.

2.5 Het Flynn Effect

James R. Flynn heeft ontdekt dat de intelligentie stijgt per generatie (Flynn, in Gleitman, Fridlund & Reisberg, 1999). Hij heeft onderzoek gedaan naar dit fenomeen in veel verschillende landen en constateerde een stijging van de IQ punten per generatie in alle landen (Flynn, in Deary, 2001). Hij heeft dit het ‘Flynn effect’ genoemd en geeft aan dat oude tests zonder hernormering, geen betrouwbaar beeld meer geven van de intelligentie. Zonder deze hernormering zouden personen per decennium ongeveer drie tot vijf IQ punten hoger uitkomen. Flynn constateerde dat de stijging van de intelligentie vooral een invloed heeft op de vloeibare intelligentie (Flynn, in Deary, 2001). Dit is juist de intelligentie die gemeten wordt door niet-verbale intelligentietests en deze tests zijn dus vooral gevoelig voor het Flynn effect. Een test kan gecorrigeerd worden voor dit Flynn effect en een dergelijke correctie is in feite ook noodzakelijk voor de SON-R. Deze test is twintig jaar oud en de normen zijn niet meer representatief. De SON-R heeft in 2004 zijn normen aangepast op het Flynn effect (Tellegen, 2005). De correctie is in 2004 vastgesteld op zes IQ punten. Echter, wanneer het Flynn effect nu berekend zou worden, dan is de correctie acht IQ punten. Dit is echter maar een globale schatting van de stijging van het IQ en een hernormering van een test zou beter op zijn plaats zijn (Tellegen, 2004). De WNV-NL is twee jaar oud en daarom is het Flynn effect niet van toepassing op deze test.

2.6 ISK

In dit onderzoek wordt er getoetst of de WNV-NL een geschikt instrument is voor een bepaalde doelgroep, namelijk leerlingen uit internationale schakelklassen. Deze klassen zijn ongeveer tien jaar geleden gevormd omdat er toen grote aantallen alleenstaande minderjarige asielzoekers (AMA's) naar Nederland kwamen. Deze jongeren moesten onderwijs krijgen, vanwege de leerplicht, maar konden niet instromen in het reguliere onderwijs. Er was sprake van een taalachterstand die eerst weggewerkt moest worden (Schipperheijn, persoonlijke communicatie, 7 juni 2010). De ISK's vormen een tussenstation waar de jongeren veel les krijgen in de Nederlandse taal om ze op hetzelfde niveau te krijgen als Nederlandse jongeren. De ISK leerlingen verblijven twee jaar lang in de internationale klas. Daarna kunnen ze instromen in het reguliere VMBO. Voor de meeste leerlingen is deze stap echter te groot, vooral vanwege de taalachterstand. Immers, om een buitenlandse jongere op het niveau te krijgen van een Nederlandse leerling, is zeven jaar onderwijs nodig (Schipperheijn, persoonlijke communicatie, 7 juni 2010). Na de twee jaar die een leerling in een ISK klas doorbrengt, hebben de leerlingen nog steeds begeleiding nodig. Dit kunnen ze krijgen in het LWOO of op het PrO.

Er zijn een aantal toelatingseisen verbonden aan het LWOO en het PrO. Voordat een leerling naar dit onderwijs mag, moet hij aan deze eisen voldoen. De eisen zijn vastgesteld door de minister van OC&W en opgenomen in de Wet op het Voortgezet Onderwijs (Cals & Marijnen, 1963). Om toegelaten te worden tot het LWOO moet een leerling een IQ score hebben tussen de 75 en de 120. Wanneer het IQ hoger is dan 90, dan kan de leerling alleen toegelaten worden als de leerachterstand verklaard kan worden op basis van een sociaal-emotionele problematiek. De sociaal-emotionele problematiek wordt vastgesteld door een orthopedagoog of gz-psycholoog aan de hand van vragenlijsten (Regeling screenings- en testinstrumenten, 2009). Er wordt ook gewerkt met een viertal vaardigheden, die onder te verdelen zijn in twee leerdomeinen. Technisch lezen en spelling zijn technische domeinen en inzichtelijk rekenen en begrijpend lezen zijn inzichtelijke domeinen. Om toegelaten te worden op het LWOO moet op twee van deze vier vaardigheden een didactische achterstand van minimaal 25% geconstateerd worden. Eén van die twee vaardigheden moet ook in het inzichtelijk domein liggen. Voor het PrO moet een leerling een IQ score hebben dat tussen de 60 en de 80 punten ligt. Daarbij moet hij een didactische achterstand van minimaal 50% hebben op twee van de vier vaardigheden (Regeling screenings- en testinstrumenten, 2009). Wanneer een leerling niet aan deze eisen voldoet, kan hij niet toegelaten worden op het LWOO of het PrO en zal de leerling in het reguliere onderwijs instromen.

Voor de toelatingseisen wordt het IQ van de leerlingen bepaald. Hiervoor wordt gebruikt gemaakt van een intelligentietest die door een bevoegde deskundige wordt afgenomen bij de leerling. Onder bevoegde deskundige wordt verstaan een onderzoeker die door de eigen beroepsgroep erkend is (Regeling screenings- en testinstrumenten, 2009). Deze test moet goed gekeurd worden door de minister van OC&W. De minister laat zich hierin adviseren door de RVC (Cals & Marijnen, 1963). De test moet een betrouwbare indicatie geven van de niet verbale intelligentie van de leerling en de test moet geschikt zijn voor de doelgroep. Om te bepalen of een test een betrouwbare indicatie van de niet verbale intelligentie geeft, neemt de RVC de beoordeling van de COTAN over (Regeling screenings- en testinstrumenten, 2009). De

RVC let vooral op de punten betrouwbaarheid, normering en begripsvaliditeit van de beoordeling (Schipperheijn, persoonlijke communicatie, 7 juni 2010). Om te bepalen of een test geschikt is voor de doelgroep, heeft de RVC een certificeringscommissie ingesteld. Deze commissie bestaat uit deskundigen uit de toets- en verwijzingspraktijk en zij geven een oordeel over de geschiktheid van een test. Vervolgens wordt een test in een categorie geplaatst die aangeeft in hoeverre een test wel of niet geschikt is (Regeling screenings- en testinstrumenten, 2009). Een overzicht van de categorieën is te vinden in bijlage 1.

3. METHODEN VAN ONDERZOEK

In dit hoofdstuk wordt de onderzoeksopzet, de steekproef en de gevolgde procedure beschreven. Tot slot volgt een beschrijving van de analysemethoden die gebruikt worden in dit onderzoek.

3.1 Onderzoeksopzet

In dit onderzoek wordt er onderzocht of de WNV-NL een geschikte vervanger is voor de SON-R. Om dit te bepalen zullen deze tests getoetst worden op hun onderlinge samenhang. Om deze samenhang te bepalen zullen beide tests afgenomen worden bij dezelfde leerlingen. Hierdoor is het mogelijk om de resultaten met elkaar te vergelijken.

3.2 Steekproef

Voor het onderzoek is gebruik gemaakt van een selecte steekproef bestaande uit 32 ISK leerlingen. De steekproef bestaat uit leerlingen die allen zijn aangemeld door het Metzo college bij de IJsselgroep om een indicatie te krijgen van hun intelligentie. Er heeft geen selectie door de onderzoekers plaatsgevonden voor de steekproef omdat er geen grote groep leerlingen beschikbaar was. Alle aangemelde leerlingen hebben aan het onderzoek deelgenomen. De leerlingen hebben allemaal minimaal een jaar onderwijs gevolgd in Nederland. Bij deze leerlingen zijn beide non-verbale intelligentietests afgenomen; de SON-R en de WNV-NL. De tests zijn afgenomen over een periode van drie weken in april van 2010. Op het Metzo college in Doetinchem zijn 13 leerlingen getest en op de dependance in Winterswijk zijn 19 leerlingen getest.

3.3 Gevolgde procedure

De ISK-leerlingen die onderzocht zijn, komen van twee verschillende locaties. De onderzoekers hebben de afnames op de verschillende locaties afgewisseld, vanwege de planning van de scholen. In de eerste week zijn de onderzoekers naar het Metzo in Doetinchem geweest. In de tweede week zijn de onderzoekers naar de dependance in Winterswijk geweest. Voor de afnames van de WNV-NL zijn de onderzoekers in de eerste week twee dagen in Doetinchem geweest. Er werden in Doetinchem per dag zeven leerlingen onderzocht. In de tweede week zijn de onderzoekers vier dagen naar Winterswijk geweest met de WNV-NL. Op drie van de vier dagen zijn er vijf leerlingen getest en op de andere dag zes leerlingen. Voor de SON-R zijn de onderzoekers vier dagen naar Doetinchem geweest en zes dagen naar Winterswijk. In de eerste week zijn ze twee dagen naar Doetinchem gegaan en in de tweede week ook twee dagen. Voor de afnames in Winterswijk zijn de onderzoekers zes dagen bezig geweest om de leerlingen te testen, verdeeld over week 2 en week 3. In die dagen werden er per dag drie SON-R afnames gedaan. Op de laatste dag van de derde week zijn de resterende twee leerlingen in Winterswijk en Doetinchem getest. Dit zijn leerlingen die door afwezigheid niet eerder getest konden worden. Van de WNV-NL zijn er meer afnames gedaan, omdat er verwarring was over de leerlingen en of ze al een jaar onderwijs hadden gevolgd. Van de leerlingen waren er drie die nog geen jaar onderwijs in Nederland hadden gevolgd. Deze drie leerlingen zijn niet getest met de SON-R. Hierdoor zijn er van de WNV-NL uiteindelijk 35 afnames gedaan en van de SON-R 32 (tabel 1). De WNV-NL gegevens van deze drie leerlingen zijn niet toegevoegd aan dit

onderzoek. Op de scholen kregen de onderzoekers een ruimte toegewezen waarin de afnames gedaan konden worden. De ruimte in Doetinchem was een rustige ruimte, wat normaal dienst doet als kantoor of vergaderruimte. In Winterswijk beschikten de onderzoekers over een kleinere ruimte, vanwege gebrek aan faciliteiten. De onderzoekers kregen door de interne begeleider een aantal leerlingen toegewezen die op die dag onderzocht konden worden. De leerlingen zelf zijn door de school op de hoogte gebracht van wat er op het programma stond. Welke leerlingen de onderzoekers toegewezen kregen, is afhankelijk van hun lestijden en of ze op school aanwezig waren. De leerlingen wisten zelf wie ze konden doorsturen naar de onderzoekers. Mocht één van de leerlingen niet komen opdagen door omstandigheden, dan kon de onderzoeker naar de betreffende contactpersoon op de school om een volgende leerling te halen. Er is getest op de scholen tijdens de lestijden van negen uur 's ochtends tot vier uur 's middags. De onderzoekers hebben ieder een eigen test bij de leerlingen afgenomen. S. Abbink heeft de SON-R afgenomen, omdat hij deze test goed kent en veel ervaring heeft met de afname van de test. Y. Oude Weernink heeft de WNV-NL afgenomen. Zij had met beide tests weinig ervaring en heeft zich verdiept in de WNV-NL. Ook omdat zij al wel ervaring heeft met het afnemen van de WISC-III en de WAIS-III.

Tabel 1
Afname van de tests

	SON-R	WNV-NL
<u>Week 1</u>		
Doetinchem	6	14
Winterswijk	0	0
<u>Week 2</u>		
Doetinchem	3	0
Winterswijk	12	21
<u>Week 3</u>		
Doetinchem	1	0
Winterswijk	10	0
Totaal	32	35

3.4 Analyses

Voor de analyses zal er gebruik worden gemaakt van een aantal bewerkingen in het programma SPSS, versie 16.0.

3.4.1 Pearson PMCC

Voor de analyse van de samenhang tussen de SON-R en de WNV-NL, zal er gebruik worden gemaakt van Pearson's productmomentcorrelatiecoëfficiënt (Pearson PMCC). Deze correlatiecoëfficiënt wordt aangeduid met het symbool r en geeft de mate aan van een lineair verband tussen twee variabelen weer (Baarda, De Goede & Van Dijkum, 2007). Deze analyse zal worden toegepast op de hele steekproef ($n=32$). De samenhang tussen de totaalscores van de SON-R en WNV-NL zal getoetst worden. Ook zal de correlatie berekend worden tussen de subtests *stripverhalen* - *plaatjes ordenen* en *analogieën* - *matrix redeneren*.

3.4.2 T-toets

De t-toets is een methode om na te gaan of de gemiddelden van twee groepen afhankelijke steekproeven op een testvariabele significant van elkaar verschillen (Baarda, De Goede & Van Dijkum, 2007). In dit onderzoek zijn de uitslagen van de steekproef afhankelijk van elkaar omdat er bij dezelfde persoon twee keer een zelfde meting is gedaan (Baarda et al., 2007). De t-toets zal toegepast worden op de totaalscores van de SON-R en de WNV-NL. Daarnaast zal de t-toets ook gebruikt worden bij de vergelijking van de gemiddelde scores tussen de subtests *stripverhalen - plaatjes ordenen* en *analogieën - matrix redeneren*.

3.4.3 Standaarddeviaties en gemiddelden

Tot slot zullen de gemiddelde scores en standaarddeviaties berekend worden van de totaalscores op de SON-R en de WNV-NL en de vier subtests *stripverhalen*, *plaatjes ordenen*, *analogieën* en *matrix redeneren*. Met deze gegevens kan bepaald worden welke test en welke onderdelen gemiddeld hoger scoort c.q. scores. De standaarddeviaties geven de mate van spreiding van de scores rond het gemiddelde aan.

4. RESULTATEN VAN HET ONDERZOEK

In dit hoofdstuk zullen de kenmerken van de steekproef besproken worden. Vervolgens zullen de resultaten van de deelvragen worden gepresenteerd. De resultaten zullen beschreven worden aan de hand van de analysemethoden en het literatuuronderzoek.

4.1 Steekproefkenmerken

De steekproef bestaat uit 32 ISK leerlingen afkomstig van het Metzo college. De leeftijd van deze leerlingen ligt tussen de 13 en 18 jaar (tabel 2). Er zijn 18 jongens en 14 meisjes getest.

Tabel 2
Leeftijd van de ISK-kinderen

<i>Leeftijd</i>	<i>Aantal</i>
13	5
14	5
15	5
16	7
17	7
18	3
Totaal	32

In tabel 3 wordt het land van herkomst van de ISK leerlingen weergegeven.

Tabel 3
Land van herkomst van de ISK-leerlingen

<i>Land</i>	<i>Aantal</i>
Afghanistan	5
China	3
Colombia	1
Frankrijk	1
Irak	4
Polen	1
Somalie	11
Thailand	4
Turkije	2
Totaal	32

4.2 Toetsing deelvragen en hoofdvraag

Om de hoofdvraag te beantwoorden zullen in deze paragraaf de deelvragen getoetst worden. Voor het beantwoorden van deelvraag één en deelvraag twee is er gebruik gemaakt van de theorie en van de kennis van deskundigen. Om de deelvragen drie en vier te toetsen zijn er enkele statische analyses gebruikt, namelijk de Pearson PMCC en de t-toets. Daarnaast zijn ook de standaarddeviaties en de gemiddelden bepaald. De resultaten van deze analyses zijn weergegeven in een aantal tabellen (4 t/m 8).

4.2.1 Uitwerking deelvraag één

Om te onderzoeken of de WNV-NL en de SON-R voldoen aan de voorwaarden die gesteld zijn door de minister van OC&W is de theorie geanalyseerd. De theorie waarvan gebruik is gemaakt, is de informatie over de tests zelf, de informatie vanuit het ministerie van OC&W, de informatie van de regionale verwijzingscommissie en de informatie van de COTAN. De minister heeft de volgende voorwaarden gesteld: de test moet gekeurd zijn door de COTAN en minimaal een voldoende hebben op de punten van beoordeling van de COTAN. Verder moet de test gekeurd zijn door de RVC en opgenomen zijn in de lijst van goedgekeurde tests van de RVC. Tot slot moet de test geschikt zijn voor de doelgroep.

Zowel de WNV-NL als de SON-R zijn door de COTAN beoordeeld (Roeloffzen, persoonlijke communicatie, 1 april 2010). In tabel 4 is de uitslag van deze beoordeling te zien. De COTAN heeft de SON-R in 1990 beoordeeld en goed bevonden op alle zeven punten. Echter, de normen van de SON-R zijn in 20 jaar niet opnieuw genormeerd en de COTAN is dan van mening dat de test wegens veroudering in principe niet meer bruikbaar is. De beoordeling van de normen wordt dan onvoldoende (Evers, Lucassen, Meijer & Sijtsma, 2010, chap. 4). De WNV-NL is in 2009 beoordeeld door de COTAN en goed bevonden op vijf van zeven punten waarop een test gekeurd wordt. De betrouwbaarheid en de criteriumvaliditeit zijn voldoende bevonden. De COTAN maakt echter bij de betrouwbaarheid van de WNV-NL de opmerking dat de beoordeling voldoende alleen van toepassing is op de totale uitslag. De betrouwbaarheid van de scores van de individuele subtests is onvoldoende.

Tabel 4
COTAN beoordelingen

Test en jaar van beoordeling	SON-R 1990	WNV-NL 2009
Uitgangspunten bij de testconstructie	Goed	Goed
Kwaliteit van het testmateriaal	Goed	Goed
Kwaliteit van de handleiding	Goed	Goed
Normen	Goed	Goed
Betrouwbaarheid	Goed	Voldoende
Begripsvaliditeit	Goed	Goed
Criteriumvaliditeit	Goed	Voldoende

Bij de beoordeling van een test door de RVC, let de certificeringscommissie in het bijzonder op drie punten van de COTAN beoordeling, namelijk de betrouwbaarheid, de normering en de begripsvaliditeit. Van beide tests is de beoordeling voldoende voor de commissie om de test goed te keuren. Echter, de COTAN beoordeling van de normen

van de SON-R zal onvoldoende worden omdat de normen van de SON-R verouderd zijn. Op basis van deze informatie heeft de RVC de WNV-NL opgenomen in de lijst van toegestane tests voor ISK onderzoeken en zal de SON-R verdwijnen uit deze lijst (Schipperheijn, persoonlijke communicatie, 7 juni 2010). De certificeringscommissie toetst ook de geschiktheid van een test voor de doelgroep. Alle deskundigen van de commissie geven een oordeel over de tests en aan de hand van deze oordelen wordt de test in een categorie geplaatst (bijlage 1). De SON-R zit in zowel categorie II als categorie III. Dit betekent dat de test bruikbaar is voor ISK onderzoeken, maar dat de normen wel verouderd zijn en dat aangeraden wordt om een andere test in te zetten. De WNV-NL zit in categorie 0 en dat betekent dat er nog onvoldoende ervaring is met de test vanuit het werkveld om er een oordeel over uit te kunnen spreken.

4.2.2 Uitwerking deelvraag twee

Om te testen wat de voor- en nadelen zijn van de praktische aspecten van de WNV-NL en de SON-R is er gebruik gemaakt van de theorie over de tests, de theorie van het Flynn effect. De praktische aspecten zijn de afnameduur, het bereik van de normen en de invloed van het Flynn effect op de normen.

De handleiding van de WNV-NL stelt dat de afnameduur gemiddeld 45 minuten bedraagt. Uit ervaring van de onderzoekers blijkt dat de afnames bij de ISK leerlingen gemiddeld 35 minuten per afname duurden. De handleiding van de SON-R stelt dat de afnameduur gemiddeld anderhalf uur is. Bij de ISK leerlingen duurt een afname echter gemiddeld langer en kan een afname uitlopen tot tweeënehalf uur. De WNV-NL kan ingezet worden bij kinderen en adolescenten van 4 tot en met 21 jaar. Met dit leeftijdsbereik kunnen de ISK leerlingen van dit onderzoek vergeleken worden met mensen van hun eigen leeftijd. De SON-R kan kinderen en adolescenten testen vanaf 5,5 jaar tot en met 17 jaar. In dit onderzoek zijn er ISK leerlingen getest van 18 jaar en deze zijn bij de SON-R vergeleken met de normgroep van 17 jaar en bij de WNV-NL met de normgroep van hun eigen leeftijd. De WNV-NL is genormeerd in 2008 en daarom is het Flynn effect nog niet van toepassing op de normen. De SON-R is genormeerd in 1985 en is daarom wel onderhevig aan het Flynn effect. De normen van de SON-R zijn in 2004 gecorrigeerd voor het Flynn effect en is de correctie vastgesteld op zes IQ punten.

4.2.3 Uitwerking deelvraag drie

Om te onderzoeken in hoeverre de uitslagen van de WNV-NL en de SON-R samenhangen is een Pearson PMCC analyse uitgevoerd met als variabelen de totaal scores van de WNV-NL en de SON-R. Hieruit blijkt dat de correlatie tussen de totaalscores van de WNV-NL en de SON-R .88 bedraagt. Dit betekent dat er een sterke samenhang is tussen de uitslagen van de twee testen. De samenhang is significant, dit blijkt uit het feit dat de p-waarde kleiner dan .01 is, namelijk .00. Een waarde van .00 betekent dat de kans dat het resultaat op toeval berust minder is dan 1 op 1000 (tabel 5).

Tabel 5
Correlatie tussen WNV-NL en SON-R

Pearson Correlatie	.88
Significantie	.00
N	32

In tabel 6 is te zien dat de WNV-NL een gemiddelde score heeft van 80,94 en de SON-R een gemiddelde score van 78,72. De standaarddeviatie van de WNV-NL is 18,00 en van de SON-R 14,37.

Tabel 6
Gemiddelden en standaarddeviaties WNV-NL en SON-R

	Gemiddelde	Standaarddeviatie
WNV-NL	80,94	18,00
SON-R	78,72	14,37

Het verschil tussen de gemiddelden bedraagt 2,22 IQ punten. Uit de t-toets komt naar voren dat dit verschil niet significant is. In tabel 7 is te zien dat de t-waarde 1.45 bedraagt en de significantie .16 is.

Tabel 7
Paired sample t-toets

	T-waarde	Vrijheidsgraden	Significantie
WNV-NL - SON-R	1.45	31	.16

4.2.4 Uitwerking deelvraag vier

Om de verwachting te toetsen in hoeverre cultuurgevoeligheid en veroudering van de test items een rol spelen in de uitslagen, zijn de correlaties berekend tussen de subtests *plaatjes ordenen - stripverhalen* en *matrix redeneren - analogieën*. Hierbij is gebruik gemaakt van de Pearson PMCC analyse. Daarnaast zijn de gemiddelde scores en de standaarddeviaties berekend. Ook is met behulp van de t-toets bepaald of de verschillen tussen de gemiddelde scores significant is. Als variabelen zijn de scores van de subtests gebruikt.

Uit de Pearson PMCC analyse komt naar voren dat de correlatie tussen *plaatjes ordenen* en *stripverhalen* .69 is. De correlatie tussen *matrix redeneren* en *analogieën* bedraagt .53 (tabel 8).

Tabel 8
Correlaties tussen de subtests stripverhalen-plaatjes ordenen en analogieën-matrix redeneren

Plaatjes ordenen / Stripverhalen		Matrix Redeneren / Analogieën	
		SON-R Stripverhalen	SON-R Analogieën
Pearson Correlatie	WNV-NL Plaatjes ordenen	.69	WNV-NL Matrix redeneren .53
Significantie		.00	.00
N		32	32

In tabel 9 is te zien wat de gemiddelde scores zijn van de subtests en de standaarddeviaties.

Tabel 9
Gemiddelden en standaarddeviaties van de subtests

	Gemiddelde	Standaarddeviatie
WNV-NL Plaatjes ordenen	89,09	17,36
WNV-NL Matrix redeneren	88,28	12,22
SON-R Stripverhalen	83,62	19,06
SON-R Analogieën	81,16	14,67

In tabel 10 zijn de resultaten te zien van de t-toets. Hieruit komt naar voren dat het verschil in gemiddelde scores tussen *matrix redeneren* en *plaatjes ordenen* niet significant is. De t-waarde is .30 en de significantie is .77. Het verschil in gemiddelde scores tussen *analogieën* en *stripverhalen* is eveneens niet significant. De t-waarde is 1.00 met een significantie van .32.

Tabel 10
Paired sample t-toets

	T-waarde	Vrijheidsgraden	Significantie
WNV-NL MR - WNV-NL PO	.30	31	.77
SON-R ANA - SON-R SV	1.00	31	.32

5. CONCLUSIES EN AANBEVELINGEN

Dit hoofdstuk zal duidelijk maken wat de antwoorden zijn op de vragen die in dit onderzoek centraal staan en wat de conclusie van de hoofdvraag is. Daarbij worden enkele bijzonderheden van het onderzoek vermeld en tot slot volgen enkele aanbevelingen.

5.1 Conclusies

In deze paragraaf zal per deelvraag een antwoord gegeven worden op de vraag en hieruit volgt de conclusie van de hoofdvraag. Tot slot worden er overige bijzonderheden met betrekking tot het onderzoek beschreven.

5.1.1 Conclusie deelvraag één

‘Voldoen de WNV-NL en de SON-R aan de voorwaarden die gesteld zijn door de minister van OC&W?’

De WNV-NL voldoet aan de voorwaarden die gesteld zijn door de minister van OC&W. Echter de SON-R voldoet niet aan de voorwaarden die gesteld zijn door de minister van OC&W. Met betrekking tot de COTAN gegevens, blijkt uit dit onderzoek dat de WNV-NL goed bevonden is. Op vijf van de zeven punten heeft de WNV-NL de beoordeling goed gekregen en twee van de zeven punten een voldoende. De SON-R is op alle punten goed bevonden, maar de normen zijn twintig jaar oud en dat betekent dat de normen de beoordeling onvoldoende krijgen. De RVC let voornamelijk op de normen bij de beoordeling van de test. Wanneer deze onvoldoende is, zal de RVC de test niet opnemen in de lijst van goedgekeurde tests. De deskundigen van de certificeringscommissie hebben een oordeel gegeven over de SON-R. Op basis daarvan is de test geschikt voor de doelgroep en bruikbaar voor onderzoeken bij ISK leerlingen, wel wordt aangeraden om een andere test te gebruiken waarvan de normen van de test niet verouderd zijn. De deskundigen van de commissie hebben onvoldoende ervaring met de WNV-NL om een oordeel te geven over de geschiktheid van de afname voor ISK onderzoeken.

5.1.2 Conclusie deelvraag twee

‘Wat zijn de voor- en nadelen van de praktische aspecten van de WNV-NL en de SON-R?’

Het eerste voordeel van de WNV-NL is dat de afnameduur gemiddeld korter is dan die van de SON-R. Uit ervaring van de onderzoekers blijkt dat de afnameduur van de WNV-NL gemiddeld 35 minuten bedraagt. Een afname van de SON-R bedraagt gemiddeld twee uur. Dit betekent dat er op een dag meer afnames van de WNV-NL gedaan kunnen worden dan van de SON-R. Een tweede voordeel van de WNV-NL is dat de test meer normgroepen heeft dan de SON-R. De WNV-NL kan personen testen vanaf 4 tot en met 21 jaar. De SON-R kan personen testen van 5,5 tot en met 17 jaar. Als derde voordeel van de WNV-NL is dat de test recentelijk genormeerd is en het Flynn-effect nog niet van invloed is op de normen. De SON-R is in 1985 genormeerd, dit betekent dat het Flynn-effect van toepassing is op de normen. In 2004 zijn de normen

gecorrigeerd voor het Flynn-effect. Op dat moment bedroeg het Flynn-effect zes IQ-punten. Wanneer het Flynn-effect nu berekend zou worden, zou deze uitkomen op een verschil van acht IQ-punten. De correctie die toegepast is op de normen van de SON-R is nu ook verouderd.

5.1.3 Conclusie deelvraag drie

‘Hoe groot is de samenhang tussen de uitslagen van de SON-R en de WNV-NL?’

Zoals verwacht, komt uit dit onderzoek naar voren dat er sprake is van een sterke correlatie tussen de totaalscores van de WNV-NL en de SON-R ($r=.88$). Een correlatie van .75 of hoger duidt op soortgenootvaliditeit en dit is bij deze tests het geval. De hoge samenhang tussen de tests geeft weer dat beide tests een betrouwbare indicatie geven van de niet-verbale intelligentie bij de ISK leerlingen.

Uit dit onderzoek blijkt verder dat de ISK leerlingen gemiddeld hoger scoren op de WNV-NL, dan op de SON-R, namelijk 2,22 IQ punten. Dit is echter geen significant verschil, wat blijkt uit de t-toets analyse: $t(31)=1.45$, $p=.16$.

5.1.4 Conclusie deelvraag vier

‘In hoeverre spelen cultuurgevoeligheid en veroudering van de test items een rol in de uitslagen?’

De correlatie tussen *stripverhalen* en *plaatjes ordenen* bedraagt .69. Dit is hoger dan de correlatie tussen *analogieën* en *matrix redeneren* ($r=.53$). Beide correlaties zijn significant ($\alpha<.00$). Hieruit is af te lijden dat de correlatie van de subtests met concrete inhoud (*stripverhalen* en *plaatjes ordenen*) dichter bij de correlatie ligt van de hele test, dan de correlatie van de subtests met abstracte inhoud (*analogieën* en *matrix redeneren*). Wanneer er sprake zou zijn van cultuurgevoelige en verouderde test items, dan is de verwachting dat de correlatie tussen de subtests met concrete inhoud lager uit zou komen dan de correlatie tussen de subtests met abstracte inhoud. Omdat dit niet het geval is, zou het erop kunnen wijzen dat de subtests weinig of geen items bevatten die cultuurgevoelig of verouderd zijn. Het zou ook kunnen zijn dat de leerlingen weinig tot geen last hebben gehad van de cultuurgevoelige en verouderde items die aanwezig zouden kunnen zijn in de subtests.

Daarnaast zien we dat de leerlingen gemiddeld hoger scoren op *plaatjes ordenen* dan op *matrix redeneren* van de WNV-NL. Dit betekent dat de leerlingen gemiddeld iets beter presteren op taken die concreet materiaal bevatten. Het verschil tussen de subtests is echter niet significant; $t(31)=.30$, $p=.77$. Hetzelfde is te zien bij de SON-R waar de leerlingen gemiddeld beter presteren bij de subtest *stripverhalen* dan op de subtest *analogieën*. Het verschil tussen deze twee subtests is groter dan het verschil tussen de subtests van de WNV-NL, maar ook hier is het verschil niet significant; $t(31)=1.00$, $p=.32$. Het blijkt dat de leerlingen beter presteren op taken die items bevatten met plaatjes die cultuurgevoelig en/of verouderd zouden kunnen zijn. Ze presteren echter niet significant beter op de subtests met concrete inhoud en daarom kan niet geconcludeerd worden of cultuurgevoeligheid en veroudering van de test items een rol spelen in de uitslagen.

5.1.5 Conclusie hoofdvraag

‘Is de WNV-NL een geschikte vervanger van de SON-R?’

Uit dit onderzoek komt naar voren dat de WNV-NL een geschikte vervanger is voor de SON-R. De WNV-NL voldoet aan de eisen die gesteld zijn door de minister van OC&W. De SON-R voldoet echter niet meer de eisen omdat de COTAN de normen van de SON-R onvoldoende zal beoordelen wegens veroudering. De WNV-NL is ook praktisch in gebruik, vooral vanwege de kortere afname duur. Daarnaast is het normeringsonderzoek recentelijk, in 2008, uitgevoerd en zijn de normen niet onderhevig aan het Flynn effect. De samenhang tussen de totaalscores van de WNV-NL en de SON-R is sterk ($r=.88$). Dit betekent dat beide tests een goede indicatie geven van de niet-verbale intelligentie. Tot slot is er geen nadelig effect geconstateerd door invloeden van cultuurgevoelige en verouderde items. Het is uit dit onderzoek niet duidelijk geworden of de subtests items bevatten die cultuurgevoelig en verouderd zijn. De subtests die deze items zouden kunnen bevatten, scoren licht hoger dan de subtests die geen cultuurgevoelige en verouderde items hebben, maar het verschil is niet significant.

5.1.6 Overige bijzonderheden

De resultaten die berekend zijn in dit onderzoek, zijn gebaseerd op scores die verkregen zijn met behulp van de normen die meegeleverd zijn met de test. Dit betekent dat bij de berekening van de SON-R scores de normen zijn gebruikt die in 2004 zijn gecorrigeerd voor het Flynn effect. Wanneer nu, in 2010, de correctie berekend zou worden, dan zou de correctie acht IQ punten in plaats van zes IQ punten bedragen. Dit zou geen verschil maken voor de correlatie, maar wel voor het berekenen van de gemiddelde scores. Echter, uit recentelijk onderzoek komt naar voren dat het Flynn effect aan het afnemen zou zijn (Tellegen, 2004). Er kan dus niet met zekerheid gezegd worden dat het IQ nog steeds met drie punten per 10 jaar stijgt.

5.2 Aandachtspunten betreffende het onderzoek

Voor dit onderzoek is gebruik gemaakt van een steekproef van 32 leerlingen. Dit is een vrij kleine steekproef en om betrouwbare resultaten te verkrijgen zou een grotere steekproef van 100 leerlingen beter zijn geweest (Baarde & De Goede, 2006). Dit zou echter veel tijd hebben gekost. Omdat de steekproef wel groter is dan 30, is het mogelijk geweest om statistische analyses uit te voeren op de steekproef. Ook zijn de correlaties significant, ondanks de kleine steekproef. Dit zou erop kunnen wijzen dat een grotere steekproef niet nodig is geweest. Daarnaast kunnen andere factoren een invloed hebben gehad op de test scores. Tellegen (2004) noemt dat de gezondheidssituatie van de proefpersoon, het moment van de afname en storende factoren (zoals een ongeschikte stoel, lawaai, slecht licht) voor ruis in de uitslagen kan zorgen. Om de invloed van ruis in een onderzoek als deze zo klein mogelijk te maken, is een grotere steekproef nodig.

In deze scriptie is regelmatig genoemd dat de gebruikte intelligentietests een indicatie geven van de niet-verbale intelligentie. Intelligentietests in het algemeen geven een zo betrouwbaar mogelijke schatting van de intelligentie van een persoon. Volgens Tellegen (2004) zijn de indicaties verre van nauwkeurig en de IQ scores kunnen zelfs

tien IQ punten naar boven en naar beneden afwijken. Belangrijke beslissingen nemen op basis van 'een IQ' is onverstandig omdat de intelligentie niet te omvatten is in één cijfer. Daarnaast is volgens Resing & Drenth (2007) juist de gekristalliseerde intelligentie een goede voorspeller voor school- en werkprestaties. In de ISK onderzoeken wordt alleen de vloeiende intelligentie gemeten en de onderzoekers vragen zich af in hoeverre de vloeiende intelligentie een goede voorspeller is voor de schoolprestaties van de ISK leerlingen.

5.3 Vervolgonderzoek

Voor een vervolgonderzoek is het interessant om de uitkomsten van deelvraag vier te onderzoeken. In deelvraag vier komt naar voren dat de leerlingen gemiddeld beter presteren op opdrachten met concrete inhoud, dan op opdrachten met een abstracte inhoud. Men zou verwachten dat juist het omgekeerde geconstateerd zou worden, omdat abstracte figuren minder gevoelig zijn voor cultuur en veroudering dan concrete afbeeldingen. Er zou onderzocht kunnen worden waarom de leerlingen gemiddeld beter presteren op de opdrachten met concrete inhoud, dan op opdrachten met abstracte inhoud.

Vanwege de kleine steekproef, zou het een mogelijkheid kunnen zijn om dit onderzoek uit te breiden door meer leerlingen te testen met de WNV-NL en de SON-R. De gegevens kunnen dan aan dit onderzoek toegevoegd kunnen worden om zo een grotere steekproef te krijgen. Echter, de SON-R zal geen optie meer zijn voor ISK onderzoeken vanwege de verouderde normen. Het is daarom niet aan te raden om dit onderzoek te vervolgen met de WNV-NL en de SON-R. Het is wel interessant om in een onderzoek de WNV-NL met andere niet-verbale intelligentietests te vergelijken. Zo komt in de nabije toekomst de SON-E 6-60 uit, die de SON-R 5,5-17 zal vervangen (De SON-E 6-60, n.d.).

Met betrekking tot het aandachtspunt over de indicatie van intelligentietests en het feit dat de gekristalliseerde intelligentie een betere voorspeller voor schoolprestaties is, zou het interessant zijn om te onderzoeken hoe het met de ISK leerlingen gaat nadat ze in het LWOO of PrO ingestroomd zijn. Heeft het ISK onderzoek een goede voorspelling gegeven van de mogelijkheden van de leerling? Of zijn er leerlingen die beter of juist slechter presteren op de opleiding dan verwacht werd aan de hand van het ISK onderzoek? Er zou onderzocht kunnen worden of het nemen van belangrijke beslissingen aan de hand van één IQ nadelige effecten heeft voor de leerling en of de vloeiende intelligentie een hoge predictieve validiteit heeft ten opzichte van de school- en werkprestaties.

5.4 Aanbevelingen

Naar aanleiding van het onderzoek kunnen een aantal aanbevelingen gegeven worden. De aanbeveling van de onderzoekers is dat de SON-R vervangen zal worden door een andere test. De normen van de SON-R zijn sterk verouderd en zullen van de COTAN een onvoldoende beoordeling krijgen. Dit betekent dat de RVC de test niet zal goedkeuren voor ISK onderzoeken. Daarmee voldoet de SON-R niet meer aan de voorwaarden die gesteld zijn door de minister van OC&W. De SON-R zal ook vervangen worden door de SON-E. Dit betekent dat de SON-R en zijn normen niet meer vernieuwd

zullen worden en dat de test waarschijnlijk in de toekomst niet meer leverbaar zal zijn.

Verder kunnen de onderzoekers de WNV-NL aanbevelen als vervanger van de SON-R. De samenhang tussen de WNV-NL en de SON-R is sterk, wat aangeeft dat de WNV-NL een betrouwbare schatting geeft van de niet-verbale intelligentie bij de doelgroep. De WNV-NL bevat nieuwe normen, is door de COTAN goed bevonden en voldoet aan de voorwaarden van de minister. Daarnaast is de afnameduur van de WNV-NL korter dan de afnameduur van de SON-R waardoor er meer afnames gedaan kunnen worden in dezelfde tijd. Met de SON-R kunnen 50 leerlingen in 25 dagen onderzocht worden, uitgaande van twee afnames per dag. Met de WNV-NL kunnen 50 leerlingen in 10 dagen getest worden, berekend over 5 afnames per dag.

Omdat de WNV-NL nog maar recentelijk op de markt is, heeft de RVC de test in categorie 0 geplaatst. Dit betekent dat de deskundigen nog te weinig ervaring hebben met de test. De onderzoekers bevelen de IJsselgroep aan dat de beoordeling van de deskundigen in de gaten wordt gehouden en in welke categorie zij de WNV-NL zullen plaatsen. Wanneer de WNV-NL door de deskundigen in categorie III of IV wordt geplaatst, is de test niet geschikt voor onderzoeken bij ISK leerlingen. Aan de hand van dit onderzoek is de verwachting van de onderzoekers dat dit echter niet zal gebeuren.

Tot slot bevelen de onderzoekers aan om voorzichtig te zijn bij de interpretatie van de test scores. Het is onverstandig om belangrijke beslissingen, zoals plaatsing van de ISK leerling in het LWOO en het PrO, te baseren op een enkel IQ getal. Uit dit onderzoek is gebleken dat de samenhang tussen de WNV-NL en SON-R sterk is. De verschillen tussen de IQ scores van de leerlingen is niet groot, maar wel aanwezig. Een leerling komt bijvoorbeeld op de WNV-NL uit op een IQ van 90 en op SON-R de op een IQ van 86. Het verschil tussen deze scores is niet groot, maar heeft wel een directe invloed op de voorwaarden van aanmelding bij het LWOO. Met een IQ van 86 mag de leerling 'gewoon' naar het LWOO, maar bij een IQ van 90 moet eerst de leerachterstand verklaard worden op basis van sociaal-emotionele problematiek. Daarnaast meet een niet-verbale intelligentietest maar een deel van de analytische intelligentie. De IQ score die verkregen wordt, geeft dus niet de algemene intelligentie weer van de leerlingen.

LITERATUURLIJST

- American Psychological Association's Task Force. (1995). *Intelligence: Knowns and Unknowns*. Geraadpleegd 15 juni, 2010, van http://www.lrainc.com/swtaboo/taboo/apa_01.html
- Baarda, D. B. & Goede, M. P. M. de. (2006). *Basisboek Methoden en Technieken*. (4th ed.). Groningen: Wolters-Noordhoff bv.
- Baarda, D. B., Goede, M. P. M. de. & Dijkum, C. J. van. (2007). *Basisboek: Statistiek met SPSS*. Groningen: Wolters-Noordhoff bv.
- Brybaert, M. (2006). *Psychologie*. Gent: Academia Press
- Cals, J. & Marijnen, V. G. M. (1963). *Wet op het voortgezet onderwijs*. Geraadpleegd 7 juni 2010, van <http://wetten.overheid.nl/BWBR0002399/Slotformulierenondertekening>
- Deary, I. J. (2001). *Intelligence: A very short introduction*. Oxford: Oxford University Press
- Deelman, B., Eling, P., Haan, E. de. & Zomeren, E. van. (2006) *Klinische neuropsychologie*. (7th ed.). Amsterdam: Uitgeverij Boom.
- DeHart, G. B., Sroufe, L. A. & Cooper, R. G. (2004). *Child development: It's nature and course*. New York: McGraw-Hill
- De SON-E 6-60: De ontwikkeling van een nieuwe versie van de SON-test*. (n.d.) Geraadpleegd op 18 augustus 2010 van <http://www.testresearch.nl/son-e/index.html>
- Driscoll, M. P. (2005). *Psychology of learning for instruction*. (3rd ed.). Boston: Pearson Education, Inc.
- Evers, A., Lucassen, W., Meijer, R. & Sijtsma, K. (2010). *COTAN Beoordelingsstelsel voor de kwaliteit van tests*. Geraadpleegd 10 augustus 2010 van <http://www.cotan.nl/website/wat-doet-het-nip/tests/beoordelingsprocedure/beoordelingsprocedure>
- Gleitman, H., Fridlund, A. J. & Reisberg, D. (1999). *Psychology*. (5th ed.). New York: W. W. Norton & Company, Inc.
- Hofstede, G. H. (1997). *Cultures and organisations: Software of the mind*. (2nd ed.). New York: McGraw-Hill
- Huizing, E. (2003). *Inleiding SPSS 11 voor windows*. Schoonhoven: Academic Service.
- Kievit, Th., Tak, J. A. & Bosch, J. D. (2002). *Handboek psychodiagnostiek voor de hulpverlening aan kinderen*. (6th ed.). Utrecht: De Tijdstroom Uitgeverij.

- Plucker, J. A. (Ed.). (2003). *Human intelligence: Historical influences, current controversies, teaching resources*. Geraadpleegd 8 juni, 2010, van <http://www.indiana.edu/~intell>
- Regeling screenings- en testinstrumenten LWOO en PrO schooljaar 2010-2011. (2009) Geraadpleegd 7 juni, 2010, van Staatscourant Web site: <https://zoek.officielebekendmakingen.nl/stcrt-2009-10924.html>
- Resing, W. & Drenth, P. (2007). *Intelligentie: Weten en meten*. (2nd ed.) Amsterdam: Uitgeverij Nieuwezijds.
- Rijksoverheid. (2010a). *Wat is de leerplicht?*. Geraadpleegd 20 augustus 2010, van <http://www.rijksoverheid.nl/onderwerpen/leerplicht/vraag-en-antwoord/wat-is-de-leerplicht.html>
- Rijksoverheid. (2010b). *Wat is leerwegondersteunend onderwijs?*. Geraadpleegd 4 juli, 2010, van <http://www.rijksoverheid.nl/onderwerpen/voortgezet-onderwijs/vraag-en-antwoord/wat-is-leerwegondersteunend-onderwijs-lwoo.html>
- Rijksoverheid. (2010c). *Wat is praktijkonderwijs?*. Geraadpleegd 4 juli, 2010, van <http://www.rijksoverheid.nl/onderwerpen/voortgezet-onderwijs/vraag-en-antwoord/wat-is-praktijkonderwijs.html>
- Roediger, H. L., Capaldi, E. D., Paris, S. G., Polivy, J., Herman, C. P. & Brysbaert, M. (2004). *Psychologie, een inleiding*. (4th ed.) Gent: Academia Press.
- Snijders, J. Th., Tellegen, P. J. & Laros, J. A. (1997) *Snijders-Oomen niet-verbale intelligentietest: SON-R 5½-17*. Lisse: Swets & Zeitlinger B.V.
- Tellegen, P. J. (2004). *De waan van 'het' IQ*. Groningen: Rijksuniversiteit Groningen, Vakgroep Persoonlijkheds- en Differentiële Psychologie.
- Tellegen, P. J. (2005). *SON-test: De Windows-versie van het computerprogramma van de SON*. Geraadpleegd 15 juli, 2010, van <http://www.testresearch.nl/sonr/sonrprgr.html>
- Tellegen, P. J., & Laros, J. A. (2004). Cultural Bias in the SON-R Test: Comparative Study of Brazilian and Dutch Children. *Psicologia*, 20. 103-111.
- Wechsler, D. & Naglieri, J. A. (2008). *Wechsler Nonverbal Scale of Ability Nederlandstalige bewerking: Technische handleiding*. Amsterdam: Pearson Assessment and Information B.V.
- WNV-NL: *Wechsler Non Verbal-NL*. (n.d.). Geraadpleegd 20 juli, 2010, van <http://www.pearson-nl.com/producten/164-wnv-nl-wechsler-non-verbal-nl.html>

Bijlagen

		Pagina
Bijlage 1	Beschrijving categorieën RVC	37
Bijlage 2	Interview met Leo Schipperheijn	38

Bijlage 1 Beschrijving categorieën RVC

Categorie 0: Er is nog onvoldoende ervaring met dit instrument in het veld en het komt nog te weinig voor in het landelijke databestand om een oordeel over de geschiktheid dan wel bruikbaarheid voor de doelgroep(en) LWOO- en/of PrO-leerlingen uit te kunnen spreken.

Categorie I: Volgens de uitslag van de enquête is dit instrument geschikt dan wel bruikbaar voor de afname bij potentiële LWOO- en/of PrO-leerlingen; de analyse van de gegevens uit het landelijke databestand ondersteunt de enquête uitslag. De normering is niet ouder dan 15 jaar.

Categorie II: Volgens de enquête uitslag is dit instrument geschikt dan wel bruikbaar voor afname bij potentiële LWOO- en/of PrO-leerlingen; ook uit het landelijke databestand blijken geen negatieve aspecten voor de beoordeling op dit punt; de normering is echter nogal verouderd (15 jaar; ingeval een test de enige is op zijn gebied zelfs ouder dan 15 jaar).

Categorie III: Meer dan de helft van de respondenten vonden dit instrument niet zomaar geschikt of bruikbaar bij potentiële LWOO- en/of PrO-leerlingen en/of de analyse van de gegevens uit het landelijke databestand geeft aan dat dat aan het gebruik van dit instrument voor de doelgroep(en) (potentiële LWOO- dan wel PrO-leerlingen) negatieve effecten kleven die van dien aard en omvang zijn dat aangeraden wordt het gebruik van een ander instrument voor het te onderzoeken criterium in overweging te nemen.

Categorie IV: 70 percent of meer van de respondenten vonden dit instrument niet geschikt of bruikbaar voor potentiële LWOO- dan wel PrO-leerlingen en/of de analyse van de gegevens uit het landelijke databestand geeft aan dat aan het gebruik van dit instrument bij de doelgroep(en) zeer negatieve effecten kleven die van dien aard en omvang zijn dat het gebruik van dit instrument bij de onderhavige doelgroep ernstig ontraden moet worden.

Bijlage 2 Interview met Leo Schipperheijn

Hieronder staan de vragen die gesteld zijn tijdens een interview met Leo Schipperheijn op 7 juni 2010.

Welke opleidingen heeft u gedaan?

Wat zijn uw werkzaamheden binnen de IJsselgroep en de RVC?

Hoe zijn de ISK's tot stand gekomen?

Waarom gaan de leerlingen naar een ISK?

Hoe lang duurt een verblijf in een ISK?

Is er een maximum leeftijd voor een ISK?

Wat zijn de voordelen van ISK's?

Waarom is een intelligentieonderzoek nodig?

Welke eisen stelt het RVC aan de intelligentietests?

Wat is LWOO en PrO?

Wat zijn de toelatingscriteria voor LWOO en PrO?