

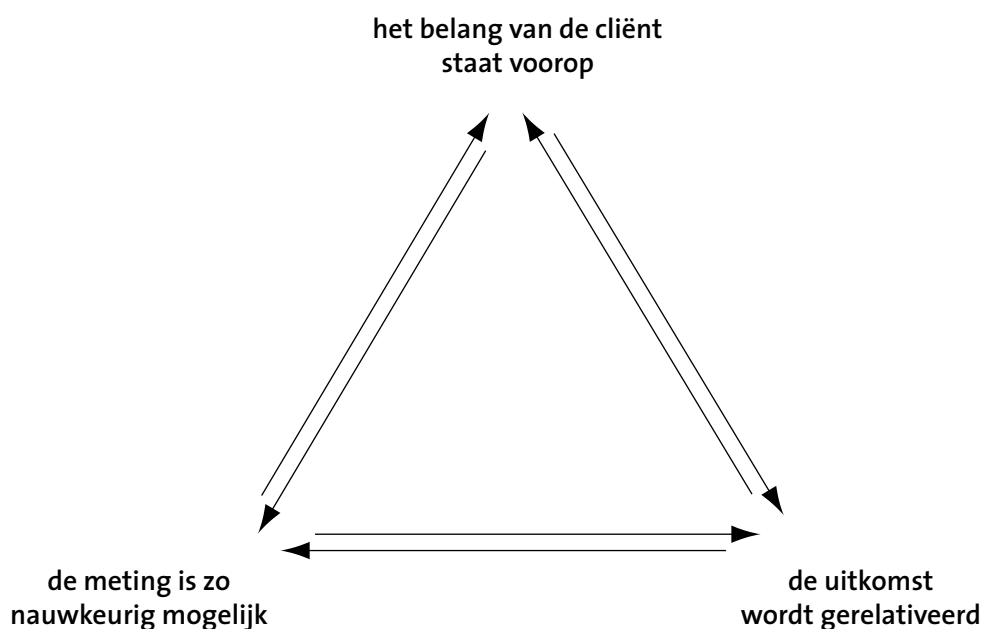
# 11 DE BETEKENIS VAN HET ONDERZOEK VOOR DE PRAKTIJK

In de voorafgaande hoofdstukken is een gedetailleerd overzicht gegeven van de resultaten van het onderzoek dat met de SON-R 6-40 is verricht. In dit hoofdstuk zullen een aantal belangrijke uitkomsten worden samengevat, waarbij de volgende vragen een centrale rol spelen:

- zijn de doelstellingen van de herziening gerealiseerd,
- geeft de test een valide meting van de intelligentie,
- voor welke personen is de test geschikt,
- hoe dient men de testresultaten te interpreteren.

## 11.1 UITGANGSPUNTEN BIJ DE CONSTRUCTIE VAN DE SON-TESTS

Bij de beoordeling van de uitkomsten van het onderzoek dienen de uitgangspunten te worden betrokken die bij de ontwikkeling van de SON-tests centraal staan. Deze uitgangspunten zijn in het kader van het symposium ter gelegenheid van het zestig jarig bestaan van de SON-tests door ons geformuleerd (Conradi, 2004) en zijn in Figuur 11.1 schematisch weergegeven. De SON-tests zijn indertijd ontstaan vanuit de sterke betrokkenheid van Nan Snijders-Oomen bij de ontwikkeling van kinderen met communicatieve handicaps. Voor ons staat als uitgangspunt nog steeds op de eerste plaats dat de test bedoeld is om het belang van de cliënt te dienen. Degene die wordt getest moet erop kunnen vertrouwen dat het diagnostisch onderzoek gericht is op zijn of haar welzijn, evenzeer als dat bij medisch onderzoek een vanzelfsprekend uitgangspunt is.



**Figuur 11.1** Uitgangspunten bij de constructie van de SON-tests

De testafname kan gezien worden als een meetprocedure waarbij op een gestandaardiseerde wijze een getalsmatige uitkomst wordt verkregen. Als men een meting verricht dan ligt het voor de hand, althans dat zou het moeten zijn, dat men probeert die meting zo nauwkeurig mogelijk te verrichten. Dat is in het belang van de cliënt.

Het streven naar nauwkeurigheid betekent echter niet dat de uitkomst verabsoluteerd moet worden. Juist het streven naar verbetering van de test maakt ook duidelijk hoe groot de beperkingen van de testuitkomst zijn in relatie tot het complexe intelligentiebegrip. Bovendien zal voor de vraagstelling die aan de testafname ten grondslag ligt intelligentie in het algemeen maar één van de relevante factoren zijn. De relativering van de testuitkomst is dan ook evenzeer in het belang van de cliënt. Daarom is het proces van validering van een test niet alleen erop gericht om aan te tonen hoe goed de test is, maar juist ook om te laten zien wat de beperkingen van de test zijn.

Het uitgangspunt voor de ontwikkeling van de SON-tests: (a) het belang van de cliënt dient voorop te staan; (b) de meting moet zo nauwkeurig mogelijk zijn en (c) de uitkomst dient gerelativeerd te worden, kan ook leidraad te zijn voor degene die de test afneemt en de uitkomsten rapporteert.

## 11.2 HET DOEL VAN DE HERZIENING

Bij de constructie van de SON-R 6-40 waren de belangrijkste doelen:

- actualiseren van de normen,
- het testmateriaal minder cultuurafhankelijk maken,
- verbeteren van het differentiatievermogen,
- geschikt maken van de test voor volwassenen,
- verkorten van de test met behoud psychometrische kwaliteiten,
- tot stand brengen van een goede aansluiting bij de SON-R 2½-7,
- doorvoeren van praktische verbeteringen.

### Actualisering van de normen

De SON-R 6-40 heeft nu normen die gebaseerd zijn op een steekproef van 1933 personen die als representatief voor Nederland en Duitsland kan worden beschouwd. De uitkomsten van beide normgroepen vertoont een grote overeenkomst hetgeen het vertrouwen in de resultaten vergroot. De aanzienlijke omvang van de steekproef betekent dat de normen met een relatief grote nauwkeurigheid berekend kunnen worden. Dit is vooral belangrijk voor de meer extreme scores. Daarnaast is het continue normeringsmodel van de SON toegepast waardoor bij de normering optimaal van alle gegevens gelijktijdig gebruik wordt gemaakt. Met het computerprogramma kunnen vervolgens de genormeerde scores op grond van de exacte leeftijd worden berekend.

### Testmateriaal minder cultuurafhankelijk maken

Het testmateriaal is beperkt doordat drie subtests uit de SON-R 5½-17, te weten Situaties, Stripverhalen en Zoekplaten, niet in de SON-R 6-40 zijn opgenomen. Hierdoor is Categorieën het enige onderdeel in de SON-R 6-40 waarbij met betekenisvol plaatjesmateriaal wordt gewerkt. Er is veel aandacht aan besteed om het plaatjesmateriaal minder cultuurafhankelijk te maken en de moeilijkheidsopbouw vergelijkbaar voor verschillende culturen. Uit de vergelijking met de Chinese normgegevens blijkt dat de relatieve moeilijkheid van de items goed met elkaar overeen komt. Dit geldt ook voor het onderzoek in Ethiopië waar in vergelijking tot Analogieën wel veel lager op Categorieën wordt gepresteerd. Allochtonen waarvan beide ouders niet in Nederland zijn geboren doen het daarentegen weer goed op dit onderdeel. Een test als Categorieën heeft altijd een culturele inhoud, maar deze is nu in veel mindere mate beperkt tot de West-Europese cultuur.

### Verbetering van het differentiatievermogen

Om de test beter geschikt te maken voor personen met beperkte capaciteiten is het differentiatievermogen van de subtests aan de 'onderkant' verbeterd door meer gemakkelijke items toe te voegen. Dit was met name van belang voor Analogieën, Categorieën en Patronen waar relatief veel gemakkelijke items zijn toegevoegd. Uit de analyse van de resultaten van het normeringsonderzoek blijkt dat de moeilijkheidsopbouw van de subtests verbeterd is en dat deze tussen de subtests beter overeenkomt. Bij het onderzoek van groepen die heel laag op de test presteren, zoals zeer moeilijk lerende kinderen in het speciaal onderwijs, blijkt dat er geen sprake is van bodemeffecten en dat er geen opeenhoping is van laagst mogelijke IQ-scores. In dit opzicht is de verbetering van het differentiatievermogen geslaagd.

Om het differentiatievermogen van de test aan de bovenkant van de verdeling te verbeteren zijn veel moeilijke items bij Mozaïeken toegevoegd en enkele moeilijke items bij Patronen. Bij Patronen is bovendien ook de maximumtijd per item verlaagd. Alhoewel bij Analogieën, Mozaïeken en Categorieën bij oudere kinderen en volwassenen niet altijd de maximale genormeerde score van 19 kan worden behaald, is de maximale score in al deze gevallen tenminste twee standaarddeviaties boven het gemiddelde. In die zin is er bij deze drie subtests geen sprake van een plafond-effect. Bij Patronen zijn de veranderingen die zijn aangebracht echter niet voldoende geweest om te voorkomen dat bij deze subtest van een plafondeffect sprake is. Vanaf veertien jaar is de hoogst mogelijke genormeerde score van Patronen lager dan zestien. Dit heeft ook gevolgen voor de IQ-score die bij de volwassenen maximaal 140 is in plaats van 145. Voor personen van gemiddeld en laag niveau discrimineert Patronen goed, echter bij volwassenen die boven gemiddeld presteren differentieert de subtest niet optimaal.

### Geslacht maken van de test voor volwassenen

Om de SON-R 6-40 voor volwassenen geschikt te maken was primair een goede normering noodzakelijk. Ofschoon het realiseren van een representatieve steekproef voor deze leeftijdscategorie veel lastiger is dan voor schoolgaande kinderen, is dit dankzij de inzet van de testleiders en de coördinator van het normeringsonderzoek wel gelukt. Het testmateriaal is in zoverre aangepast dat meer moeilijke items zijn toegevoegd en dat in de plaatjes van Categorieën meer volwassenen zijn afgebeeld. Naar de indruk van de testleiders zijn de afnames goed verlopen. Door de adaptieve procedure kunnen volwassenen die goed presteren veel gemakkelijke items overslaan. Daardoor krijgt men niet het gevoel een 'kinderachtige' test te maken. In alle gevallen werd de samenwerking als positief beoordeeld en de motivatie en het begrip van de instructies werden bij vrijwel alle volwassenen positief beoordeeld. Wel werd bij 12% van de volwassenen aangegeven dat de concentratie matig of wisselend was.

Er is een redelijke samenhang van .41 tussen het zelfoordeel van de volwassenen over hun intelligentie en het SON-IQ. De samenhang met het (gevolgde) onderwijsniveau is .64. De hoogste correlatie met een andere intelligentietest is bij de volwassenen gevonden. Deze correlatie bedroeg .83 en heeft betrekking op de samenhang van de SON-R 6-40 met de WAIS-III. De betrouwbaarheid van de SON-R 6-40 is bij de volwassenen .96 en de generaliseerbaarheid .85 hetgeen eveneens zeer bevredigend is.

### Verkorten van de test met behoud psychometrische kwaliteiten

Op grond van analyse van de resultaten van het normeringsonderzoek van de SON-R 5½-17 was duidelijk dat een beperking van het aantal subtests van zeven naar vier mogelijk was zonder dat dit ten koste hoefde te gaan van de validiteit. Op inhoudelijke en empirische gronden is hierbij gekozen voor een iets andere selectie dan de officiële verkorte vorm van de SON-R 5½-17. Ten opzichte van de volledige SON-R 5½-17 zou de betrouwbaarheid enigszins afnemen van .93 naar .91 en de genera-

liseerbaarheid zou afnemen van .86 naar .79. Door bij de vier geselecteerde subtests het aantal items uit te breiden en de moeilijkheidsopbouw te verbeteren zou de betrouwbaarheid van de subtests en de onderlinge correlaties tussen de subtests toenemen waardoor zowel de betrouwbaarheid als de generaliseerbaarheid van de totaalscore omhoog zouden gaan. Deze doelstelling is gerealiseerd: de betrouwbaarheid van het SON-IQ van de SON-R 6-40 is .95 en de generaliseerbaarheid gemiddeld .84. Of de betrouwbaarheid ten opzichte van de SON-R 5½-17 daadwerkelijk is gestegen is moeilijk te beoordelen omdat de adaptieve procedure het moeilijk maakt de betrouwbaarheid zuiver te schatten. Dit probleem geldt niet voor de generaliseerbaarheid. Deze is duidelijk toegenomen door de aanpassing van de subtests en is maar weinig lager dan van de volledig afgenomen SON-R 5½-17. Wat deze psychometrische kenmerken betreft is de overeenkomst tussen de SON-R 6-40 en de SON-R 5½-17 groot. De evaluatie van de validiteit zal in paragraaf 11.3 aan de orde komen.

De gemiddelde afnameduur van de SON-R 6-40 is 53 minuten en dit is aanzienlijk korter dan de afnameduur van de SON-R 5½-17 die 79 minuten bedroeg (beide gebaseerd op de som van de afnameduur van de subtests). De afnameduur is dus met een derde verminderd. Dit heeft praktische voordelen en maakt het beter mogelijk om desgewenst de SON-R 6-40 in combinatie met andere tests af te nemen.

### Aansluiting van de SON-R 6-40 bij de SON-R 2½-7

Bij de eerste twee versies van de SON-tests was er nog geen onderscheid tussen een test voor de jongere en een test voor de oudere kinderen. Bij de constructie van de K-SON en de SSON is dit onderscheid voor het eerst aangebracht. Tussen de K-SON en de SSON bestonden echter inhoudelijk grote verschillen, alsmede verschillen in de wijze van afname van beide tests (zie paragraaf 1.2). Bij de constructie van de SON-R 6-40 is er naar gestreefd de test goed te laten aansluiten bij de SON-R 2½-7 die in 1998 is verschenen. Door het toevoegen van gemakkelijke items bij de SON-R 6-40, is de overeenkomst tussen de moeilijke items van de SON-R 2½-7 en de gemakkelijke items van de overeenkomstige subtests van de SON-R 6-40 toegenomen. Bij beide tests wordt een adaptieve procedure gehanteerd en wordt feedback gegeven. Bij de SON-R 6-40 blijft deze feedback echter beperkt tot het aangeven of de oplossing juist of onjuist is. Beide tests werken ook met gedifferentieerde normen die met het computerprogramma berekend kunnen worden.

Zowel wat materiaal als procedures betreft, zijn er tussen de beide tests grote overeenkomsten. Er is nog geen onderzoek gedaan naar de correlatie tussen de SON-R 6-40 en de SON-R 2½-7 maar voor de SON-R 5½-17 was de correlatie met de SON-R 2½-7 gelijk aan .76.

Sattler (1992) noemt het minder wenselijk wanneer tests zoals de WPPSI en de WISC elkaar inhoudelijk overlappen omdat de tests dan niet meer als onafhankelijke tests korte tijd na elkaar kunnen worden afgenomen. De reden om de leeftijdsnormen te laten overlappen is bij de SON-tests echter niet om een hertest binnen korte tijd mogelijk te maken, maar om een keuzemogelijkheid te geven die optimaal is gezien de leeftijd, vaardigheid en specifieke problemen van de cliënt.

Een belangrijk punt waarop de aansluiting van de SON-R 2½-7 met de SON-R 6-40 is verbeterd, betreft de presentatie van de genormeerde uitkomsten. Bij beide tests hebben de subtestscores nu een gemiddelde van 10 en een standaarddeviatie van 3 en heeft het SON-IQ een gemiddelde van 100 en een standaarddeviatie van 15. Bij de SON-R 5½-17 was de normering gebaseerd op standaardisatie van de 'ware' scores en daardoor moeilijk te vergelijken met de SON-R 2½-7.

### Praktische verbeteringen

Het opbergsysteem van de test is verbeterd waardoor de verschillende onderdelen in de koffer en de deksel direct zichtbaar zijn en voor gebruik gereed. Een overzicht van de inhoud van de koffer wordt in bijlage C gegeven. Het formaat van de testboekjes is aangepast aan de vorm van de opgaven. Het

scoreformulier is ook wat kleiner gemaakt zodat het op de tafel minder ruimte in beslag neemt. De normtabellen en de instructies zijn als aparte boekjes uitgegeven zodat de gebruiker van de test alleen datgene bij zich hoeft te hebben dat strikt nodig is. Onze ervaring is dat het materiaal aantrekkelijk is voor cliënten en dat de opgaven en de instructies duidelijk zijn. Het materiaal is gemakkelijk hanteerbaar en duurzaam in gebruik.

### 11.3 DE VALIDITEIT VAN DE TEST

De vraag naar de validiteit van de test is primair de vraag of de SON-R 6-40 een goede en bruikbare maat is voor intelligentie. Voor een niet-verbale test is deze vraag extra belangrijk omdat het niet-verbale karakter beperkingen oplegt aan de wijze waarop intelligentie kan worden gemeten. De vraag is echter ook moeilijk te beantwoorden omdat intelligentie geen nauwkeurig gedefinieerd en afgebakend begrip is. Met het onderzoek dat met de SON-R 6-40 is uitgevoerd kan dan ook geen eenduidig antwoord worden gegeven over 'de' validiteit, wel geeft het onderzoek meer inzicht in de positieve kanten van de test en van beperkingen waarbij men met de interpretatie van de uitkomsten rekening moet houden. Bij de bespreking van de validiteit maken we een onderscheid tussen de inhoudelijke aspecten van de test, de samenhang met andere indicatoren van intelligentie en de samenhang met andere cognitieve en niet-cognitieve variabelen.

#### Inhoudelijke aspecten

De subtests van de SON-R 5½-17 zijn gericht op het oplossen van problemen waarbij een beroep wordt gedaan op ruimtelijk inzicht en het redeneervermogen. De prestaties zijn niet zozeer afhankelijk van verworven kennis, maar van het vermogen methoden en regels te ontdekken en deze toe te passen op nieuwe, en geleidelijk meer complexe situaties en materiaal. De SON-R 6-40 sluit hierbij aan bij een definitie van intelligentie als probleemoplossend vermogen en het vermogen te leren, en legt daarbij de nadruk op *fluid intelligence* en niet zozeer op *crystallized intelligence*. In dit opzicht is er een overeenkomst tussen de SON-tests en tests voor leerpotentieel (Tellegen & Laros, 1993a). Dit betekent echter niet dat ervaringen die iemand opdoet, niet van invloed zullen zijn op zijn of haar mogelijkheden de opgaven te kunnen oplossen.

De prestaties op de test nemen aanvankelijk sterk toe met de leeftijd. In de leeftijd van zes tot en met elf jaar wordt ruim 40% van de variantie van de ruwe totaalscore door leeftijd verklaard. In de leeftijd van zestien tot en met eenentwintig jaar is dit nog maar twee procent en bij de volwassenen zijn de prestaties stabiel. Dit betekent dat met name bij de jongere leeftijden gedifferentieerde normen noodzakelijk zijn.

De correlaties tussen de genormeerde subtestscores van de SON-R 6-40 zijn gemiddeld ,56 en nemen aanvankelijk toe om bij de volwassenen weer iets af te nemen. De gemeenschappelijke variantie van de subtests wordt vrijwel geheel door één factor bepaald. Op inhoudelijke gronden kan wel een onderscheid worden gemaakt tussen de meer ruimtelijke, visueel-motorische, performale tests (Mozaïeken en Patronen) en de tests gericht op abstract redeneren (Categorieën en Analogieën). Bij deze laatste subtests moet de goede oplossing beredeneerd worden en moet uit een aantal alternatieven gekozen worden, terwijl bij de performale subtests de oplossing geconstrueerd moet worden.

#### Congruente validiteit

De samenhang met andere indicatoren van intelligentie is onderzocht door middel van de overeenkomst met oordelen van andere personen, en door onderzoek naar de samenhang met de prestatie op andere intelligentietests.

**Tabel 11.1** Absolute verschillen tussen de scores op twee vergelijkbaar genormeerde intelligentie-tests waarbij de correlatie .80 is

absoluut verschil IQ	pct.	absoluut verschil IQ	pct.	absoluut verschil IQ	pct.
0 - 4	36%	0-4	36%	0-4	36%
5 - 9	32%	5-9	32%	5 en meer	64%
10 - 14	20%	10 en meer	32%	10 en meer	32%
15 - 19	8%	15 en meer	12%		
20 en meer	4%	20 en meer	4%		

Met het oordeel met betrekking tot intelligentie van ouders en leerkrachten in het basis- en voortgezet onderwijs en met zelfoordeel van volwassenen, correleert het het SON-IQ gemiddeld .42. Met betrekking tot het oordeel over taalvaardigheid is de gemiddelde correlatie .32.

De correlaties van de SON-R 6-40 met verbale algemene intelligentietests (WISC-III/IV, WAIS-III en NIO) zijn hoog: gemiddeld .80 met een range van .79 tot .83. De correlatie met de niet-verbale WNV is .77. Met het verbale deel van de algemene intelligentietests is de correlatie gemiddeld .69 en met het performale deel gemiddeld .79.

De hoge correlaties met algemene intelligentietests ondersteunen de aanname dat de SON-tests algemene intelligentie meten, waarbij de afname echter geheel niet-verbaal kan gebeuren. De SON-tests onderscheiden zich van het performale deel van algemene intelligentietests omdat daar minder het redeneervermogen wordt onderzocht en omdat bij deze performale tests de instructies wel verbaal zijn.

Het is belangrijk dat men zich realiseert dat wanneer intelligentietests hoog met elkaar correleren toch grote verschillen tussen de IQ-scores op kunnen treden, ook wanneer de verdeling van de genormeerde scores hetzelfde is. In Tabel 11.1 wordt een overzicht gegeven van de absolute verschillen tussen de IQ-scores op twee intelligentietests met een correlatie van .80. Hierbij wordt aangenomen dat de verschillen normaal zijn verdeeld. Gegeven een correlatie van .80 en genormeerde scores met een standaarddeviatie van 15, is de standaarddeviatie van de verschillen 9.5. Uit de berekening blijkt dat in ongeveer een derde van de gevallen het verschil redelijk beperkt is en minder dan vijf IQ-punten bedraagt. In ongeveer een derde van de gevallen is het verschil vijf tot tien punten hetgeen al belangrijke consequenties kan hebben. En in een derde van de gevallen is het verschil tien IQ-punten of meer. Bij vier op de honderd personen zal de het verschil in IQ tussen de ene en de andere test zelfs twintig punten of meer bedragen. Dergelijke verschillen kunnen in de praktijk grote gevolgen hebben.

Verschillen tussen de uitkomsten op vergelijkbare intelligentietests zijn groter dan op grond van de betrouwbaarheid van de tests te verwachten is. Hiervoor zijn verschillende oorzaken aan te wijzen:

- inhoudelijke en procedurele verschillen tussen de tests,
- wisselvalligheid van prestaties,
- stabiele veranderingen,
- onnauwkeurigheid van de normen.

### *Verschillen in inhoud*

Tussen de SON-R 6-40 en andere intelligentietests kunnen inhoudelijk grote verschillen bestaan. Specifiek verbale testonderdelen ontbreken in de SON evenals geheugentests en tests waarbij een serie handelingen moet worden nagebootst, zoals de sequentiële subtests van de K-ABC (Kaufman

& Kaufman, 1983). In een klein onderzoek bij 23 kinderen dat in Duitsland is uitgevoerd bleek het SON-IQ met de sequentiële schaal de laagste correlatie te hebben (Haarmann, 2010). De subtests van de SON-R 6-40 hebben ook geen snelheidskarakter waarbij simpele taken zo snel mogelijk moeten worden uitgevoerd. Met de *processing speed index* van de WISC-IV correleert de SON-R 6-40 dan ook relatief laag. Bij de WAIS-III hebben Symbool Substitutie en Symbool Zoeken de laagste correlatie met het SON-IQ. De samenhang met maten voor verbale intelligentie is echter hoog, mede doordat in het verbale deel van algemene intelligentietests onderdelen zijn opgenomen gericht op het redeneervermogen, zoals bijvoorbeeld tests voor verbale analogieën die een beroep doen op dezelfde vaardigheden als de redeneertaken van de SON-R 6-40.

Naast de verschillen in inhoud van de opgaven, zijn er ook verschillen in de wijze van afname die van invloed zijn, zoals de adaptieve procedure en feedback die bij de SON-R 6-40 gegeven wordt en de mogelijkheid voor de cliënt om niet te hoeven spreken. Dit kan verklaren waarom zeer moeilijk lerende kinderen bij de SON-R 6-40 zoveel beter presteren dan bij de WISC-III.

### *Wisselvalligheid van prestaties*

De wisselvalligheid van testprestaties is mede een oorzaak van verschillen tussen testcores. De her-testcorrelatie van de SON-R 6-40 is bij een interval van ongeveer vier maanden .92. Dit is niet aanzienlijk lager dan de betrouwbaarheid van de test die .95 is, maar draagt toch bij tot grotere verschillen. Het effect is echter beperkt in verhouding tot de grote mate van instabiliteit die bij de SON-R 2½-7 bij zeer jonge kinderen gevonden wordt. Op jonge leeftijd is er ook veel vaker sprake van problemen met concentratie en motivatie bij de afname van de test waardoor men bij jonge kinderen meer rekening moet houden met een zekere mate van wisselvalligheid in de prestaties. Binnen één en dezelfde testafname kan eveneens van wisselvalligheid in de prestaties sprake zijn die niet hoeft samen te hangen met de kenmerken van de subtests. Hierbij kan men denken aan factoren als vermoeidheid in de loop van de afname, fysiek ongemak, of motivatie die toeneemt wanneer de cliënt zich in de testsituatie prettiger gaat voelen.

### *Stabiele veranderingen*

Veranderingen in vaardigheid kunnen ook een meer stabiel karakter hebben en tot verschillen in scores leiden wanneer er een relatief groot interval tussen de testafnames bestaat. Het tempo waarin personen zich ontwikkelen is niet voor iedereen gelijk en zal fluctuaties vertonen. Diverse factoren hebben invloed op de cognitieve ontwikkeling en grote veranderingen in de omstandigheden waarin men opgroeit en ingrijpende gebeurtenissen, kunnen deze ontwikkeling vertragen of juist belemmeringen in de ontwikkeling opheffen. In verschillende correlationele onderzoeken met de SON-R 6-40 bedroeg de periode tussen de afname van de tests meer dan een jaar en verschillen in het verloop van de ontwikkeling kunnen leiden tot grotere discrepanties tussen de testcores.

### *Onnauwkeurigheid normen*

De onnauwkeurigheid van normtabellen kan ook een bron zijn van verschillen tussen scores. Bij tests met te grove tabellen zijn de scores rond de overgang van een leeftijdsgroep onnauwkeurig. Bij de vergelijking van testcores kan een gebrekkige overeenkomst tussen normen, als gevolg van veroudering van normen of doordat de normgroepen niet goed vergelijkbaar zijn, eveneens tot grote verschillen tussen scores leiden.

### *Conclusies*

Tezamen met de beperkingen die met de betrouwbaarheid samenhangen, lijken de specifieke inhoudelijke kenmerken van de SON-R 6-40, en in mindere mate instabiliteit van testprestaties, de belangrijkste oorzaken van de verschillen in scores tussen de SON-R 6-40 en andere intelligentietests.

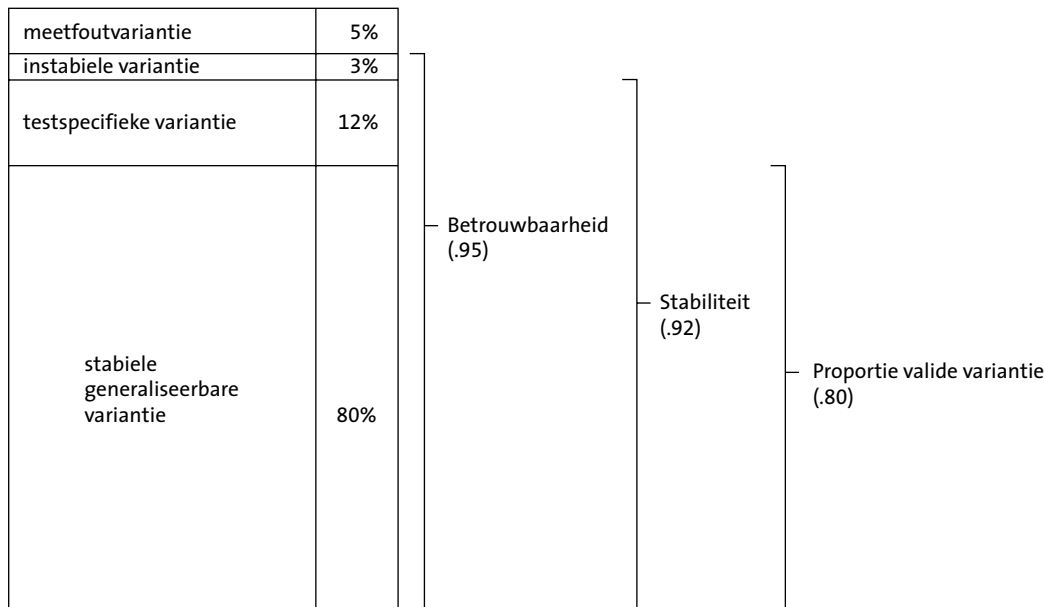


Op grond van het onderzoek dat in hoofdstuk 8 is gepresenteerd, is het realistisch om voor de correlatie van de SON-R 6-40 met andere algemene intelligentietests uit te gaan van een waarde van .80 indien het interval tussen de testafnames niet langer is dan een jaar. Op grond van deze gemiddelde correlatie met andere tests en op grond van de betrouwbaarheid en de stabiliteit, kan de variantie van het SON-IQ als volgt omschreven worden (zie ook Figuur 11.2):

- 5% meetfoutvariantie,
- 3% betrouwbare instabiele variantie,
- 12% betrouwbare testspecifieke variantie,
- 80% stabiele, betrouwbare en naar andere tests generaliseerbare variantie.

De laatste component, de variantie die de SON-R 6-40 gemeenschappelijk heeft met andere intelligentietests die op een ander moment kunnen zijn afgenomen, is het meest relevant voor de beoordeling van de intelligentie en de waarde .80 kan beschouwd worden als een indicatie van de validiteit. De proportie testspecifieke variantie zal natuurlijk afhangen van de mate waarin criteriumtests inhoudelijk en procedureel met de SON-R 6-40 overeenkomen. De validiteit wordt hier gebaseerd op de gemiddelde correlatie met een verbale algemene intelligentietest. De generaliseerbaarheidcoëfficiënt van de SON-R 6-40 is ook een validiteitsmaat die echter de verwachte waarde aangeeft van de correlatie met een soortgelijke niet-verbale intelligentietest. Gemiddeld is deze generaliseerbaarheidcoëfficiënt .84. Het blijkt dus dat de correlaties met algemene intelligentietests niet veel lager zijn dan de verwachte waarde van de correlatie van de SON-R 6-40 met een soortgelijke niet-verbale intelligentietest.

Wanneer we de scores op de SON-R 6-40 zouden kunnen correleren met de 'ideale' score, gebaseerd op een groot aantal andere algemene intelligentietests, op verschillende momenten afgenomen, dan zou de correlatie gelijk zijn aan de wortel van .80 en dit betekent dat de validiteitscoëfficiënt van de SON-R 6-40 gelijk is aan .89.



**Figuur 11.2** De componenten van de variantie van de SON-R 6-40



### Begripsvaliditeit

Bij de ontwikkeling van de eerste SON-test was het streven de eenzijdigheid van de indertijd gebruikte performale tests te doorbreken en in de niet-verbale test ook taken op te nemen die gerelateerd zijn aan abstract en concreet redeneren, taken die in meer traditionele tests vaak een verbale vorm hebben. Het doel bij de SON-tests is zo goed mogelijk algemene intelligentie te meten binnen de beperkingen die een niet-verbale afname stelt. De uitkomst van de factoranalyse van de subtests van de SON-R 6-40 ondersteunt niet een duidelijk onderscheid naar performale subtests (Mozaïeken en Patronen) en redeneertests (Categorieën en Analogieën). Het overgrote deel van de gemeenschappelijke variantie van de subtests is op één algemene factor terug te voeren. Dit is het gevolg van het feit dat Analogieën zowel een sterk ruimtelijke als een sterke redeneercomponent heeft waardoor het geen cluster met Categorieën vormt. Uit de factoranalyse komt wel een verschil naar voren tussen Categorieën enerzijds en de twee ruimtelijke subtests anderzijds, maar de betekenis van dit onderscheid is beperkt. Categorieën blijkt in verschillende gevallen relatief hoge correlaties te hebben met taalontwikkeling en verbale intelligentie en dit komt overeen met de bevindingen dat het vermogen tot abstract redeneren sterk verbonden is met de taalontwikkeling.

### Geheugen

Bij de SON-R 6-40 ontbreken, evenals bij de SON-R 2½-7, typische geheugentests. De correlatie van het SON-IQ van de SON-R 2½-7 met de TOMAL (Reynolds & Bigler, 1994) was .45 en ook de correlaties met andere geheugentests waren matig positief. Deze relatief lage samenhang pleit ervoor intelligentie en geheugen afzonderlijk te onderzoeken: het opnemen van enkele geheugen-subtests in een intelligentietest is te beperkt om een valide uitspraak over het geheugen te doen, mede omdat tests voor het lange termijn geheugen ontbreken.

### Visuele perceptie

In de SON-R 6-40 is het aspect van visuele perceptie sterk vertegenwoordigd. Vooral de performale subtests van de SON-R 6-40 hangen sterk met visuele waarneming samen. Waarneming is hier echter niet een 'passief' zien, maar houdt in het structureren, beoordelen en vergelijken van visuele informatie.

### Motorische vaardigheden

De performale subtests van de SON-R 6-40 doen een beroep op visueel-motorische vaardigheden. Bij onderzoek met de SON-R 2½-7 in het basisonderwijs was de correlatie van het SON-IQ met het oordeel van de leerkracht over de motorische ontwikkeling laag (.24). Bij de kinderen met een ontwikkelingsachterstand en bij kinderen met een taal/spraak/gehoorstoornis waren deze correlaties echter hoger (.46 en .32).

### Verbale vaardigheden

De samenhang tussen de SON-R 6-40 en verbale intelligentie en taalvaardigheid is belangrijk om te kunnen beoordelen in welke mate het intelligentiedomein door het niet-verbale karakter van de test beperkt wordt. Verbale intelligentie is echter geen duidelijk omschreven begrip. Op grond van factoranalytisch onderzoek wordt door Kaufman (1975) bij de verbale schaal van de WISC-R een onderscheid gemaakt tussen de factor *Verbal comprehension* en *Freedom from Distractibility*. Bij de subtests van de factor *Verbal comprehension* zijn er onderling weer grote verschillen tussen de subtests, zoals bijvoorbeeld tussen Overeenkomsten, een verbale subtest voor abstract redeneren, en Woordenschat, een test gericht op verbale kennis. De vaardigheden die voor de subtest Overeenkomsten nodig zijn, behoren tot het intelligentiedomein dat de SON-R 6-40 beoogt te meten; de prestaties op een subtest als Woordenschat zijn zo sterk afhankelijk van de omstandigheden waar-

onder iemand opgroeit, dat een niet-verbale variant van deze test niet in de SON-R 6-40 zou worden opgenomen. Bij de K-ABC worden dergelijke op kennis gerichte onderdelen apart gescoord en niet bij de berekening van de mentale ontwikkelingsindex betrokken.

Dat het onderscheid tussen intelligentie, verbale intelligentie en taalvaardigheid niet scherp te trekken is, blijkt ook uit de correlaties van het SON-IQ met beoordelingen van de intelligentie en de taalvaardigheid. Bij de kinderen in het basisonderwijs is de correlatie met het oordeel over de intelligentie .53 en de correlatie met het oordeel over de taalvaardigheid is met .45 ook aanzienlijk. Bij de beoordelingen door ouders zijn de correlaties met taalvaardigheid .10 lager dan met intelligentie. In de groep zeer moeilijk lerende kinderen is er een gering verschil, de correlatie met het oordeel van de leerkracht over intelligentie is .42 en de correlatie met het oordeel over taalvaardigheid is .38.

De correlaties van het SON-IQ met het verbale gedeelte van intelligentietests zijn gemiddeld .69. In aanmerking genomen dat de SON-R 6-40 geheel kan worden afgenomen zonder hierbij taal te gebruiken, zijn dit aanzienlijke verbanden. Alle vier de subtests van de SON-R 6-40 dragen aan deze correlaties bij. Tussen de subtests varieert de correlatie met de verbale schalen van .52 tot .57.

Bij de groep kinderen met communicatieve beperkingen zijn de verschillen tussen de correlatie met het verbale en het performale deel van de algemene intelligentietests niet groot: bij de WISC-III zijn de correlaties .58 en .55 en bij de NIO .62 en .71. Met de woordbegrijptest van de PPVT-III correleert de SON-R 6-40 echter vrijwel niet (.22).

In het reguliere basisonderwijs zijn de correlaties van het SON-IQ met rekenen (.53), taal (.42) en begrijpend lezen (.44) het hoogst. De samenhang met toetsen van het leerlingvolgsysteem is voor rekenen .57 en voor begrijpend lezen .55. Voor de Cito-entreetoets en de Cito-eindtoets is de correlatie met rekenen .64 en met taal .62. In de groep kinderen met communicatieve beperkingen is de correlatie met rapportcijfers en Cito-scores voor wiskunde en rekenen .39 en voor Nederlands en begrijpend lezen .46. Dit betekent dat de predictieve validiteit van de SON-R 6-40 met betrekking tot rekenen en taal hoog is en onderling weinig verschilt.

### *Sociaal-culturele verschillen*

In vergelijking tot de WNV hangt het SON-IQ minder sterk met het opleidingsniveau van de ouders samen. Vooral in de categorie met een laag opleidingsniveau is het SON-IQ hoger. Uit eerder onderzoek kwam ook naar voren dat het SON-IQ minder sterk met het SES-niveau samenhangt dan taaltests en het verbale deel van algemene intelligentietests. In vergelijking tot de meeste andere tests is er bij de SON-R 6-40 sprake van minder grote verschillen tussen autochtone en allochtone kinderen. Bij de WISC-III is de achterstand van allochtonen zelfs twee keer zo groot als bij de SON-R 6-40.

### *Conclusies*

Op grond van deze uitkomsten kan geconcludeerd worden dat het intelligentiebegrip dat met de SON-R 6-40 gemeten wordt, in grote mate overeenkomt met wat onder algemene intelligentie wordt verstaan. Hierbij ligt bij de SON-R 6-40 de nadruk op het vermogen tot abstract redeneren en ruimtelijk inzicht en visueel-motorische, perceptieve vaardigheden. In de indeling van Carroll (1993) komt dit overeen met de factoren *Fluid Intelligence* en *Broad Visual Perception*. Geheugen, kennis, en taalvaardigheid hebben wel een indirecte relatie met de prestaties, maar de meting is daar niet op gebaseerd. De test is minder dan verbale tests van sociaal-culturele factoren afhankelijk en kan het best gekenschetst worden als niet-verbale algemene intelligentietest waarbij de nadruk ligt op *fluid intelligence* en *visual perception*.

### Predictieve validiteit

Voor een intelligentietest is de samenhang met het niveau van onderwijs en met schoolprestaties een belangrijk aspect van de predictieve validiteit. Gemiddeld is de correlatie met toetsen van het leerlingvolgsysteem, de Cito-entreetoets en de Cito-eindtoets .67. De multiële correlatie met drie indicatoren van onderwijssucces (onderwijstype, schoolloopbaan en rapportcijfers) is in het basisonderwijs .64 en bij het voortgezet onderwijs .72. Dit zijn aanzienlijke verbanden. Het stelt de aanname ter discussie of een niet-verbale test zoals de SON-R 6-40 minder goed in staat zou zijn om schoolprestaties te voorspellen dan een algemene intelligentietest. Zo zijn in het basisonderwijs de correlaties van het SON-IQ met de rapportcijfers voor Rekenen en Taal .53 en .42 en voor de WISC-III zijn deze correlaties .51 en .44 (Wechsler, 2005a).

## 11.4 DE DOELGROEPEN

De SON-test heeft zich in de loop van de tijd ontwikkeld van een intelligentietest voor dove kinderen tot een algemeen bruikbare niet-verbale intelligentietest die in het bijzonder geschikt is voor kinderen met communicatieve handicaps, zoals kinderen met taal/spraak- en gehoorproblemen, kinderen met dyslexie en autistische kinderen en kinderen die anderstalig of tweetalig zijn opgegroeid. Daarnaast is de test eveneens goed bruikbaar voor kinderen die moeilijk testbaar zijn, en voor moeilijk lerende kinderen en kinderen met een ontwikkelingsachterstand. Aangezien bij de normering ook referentieleeftijden berekend worden, kan de test gebruikt worden voor oudere zwakbegaafde en verstandelijk gehandicapte kinderen en volwassenen. Door de uitbreiding van de normering is de SON-R 6-40 nu ook geschikt geworden voor volwassenen met achterstanden en taalproblemen, maar evenzeer voor normaal functionerende ouderen en hoogbegaafden. Voor bepaalde categorieën personen is de test echter niet of minder goed bruikbaar. Het gaat hierbij om personen met visuele handicaps en personen met ernstige motorische handicaps.

### Communicatieve handicaps

Uit het onderzoek bij doven met de SON-R 2½-7 en met de SON-R 5½-17, waarbij vrijwel de gehele populatie dove kinderen in de leeftijdsrange van de test is onderzocht (zie Snijders, Tellegen & Laros, 1988, hoofdstuk 9) bleek dat de achterstand van dove kinderen die niet meervoudig gehandicapt zijn ten opzichte van de horende populatie ongeveer drie IQ-punten bedraagt. De achterstand werd met name veroorzaakt door de lagere prestaties op Analogieën en Categorieën, subtests die nu in de SON-R 6-40 relatief een grote invloed hebben. Uit het onderzoek met de SON-R 5½-17 bleek dat daardoor het verschil voor de doven één IQ-punt groter zou worden zodat het gemiddeld SON-IQ van de niet meervoudig gehandicapte doven bij de SON-R 6-40 naar verwachting ongeveer 96 zal zijn. Dat doven meer moeite hebben met abstract redeneren is niet verwonderlijk aangezien de verbale ontwikkeling hierbij zo'n sterke rol speelt.

De gemiddelde score van de kinderen met een taal/spraakstoornis en/of een gehoorstoornis is 93 bij de SON-R 6-40 en bij de SON-R 2½-7 ligt dit rond de 90. Deze groepen kinderen zijn echter niet goed vergelijkbaar met de groep dove kinderen. Enerzijds zijn er in deze groep wel kinderen die meervoudig gehandicapt zijn, anderzijds zijn kinderen die met een taal/spraak/gehoorstoornis in het reguliere onderwijs goed functioneren niet in de onderzoeksgroep opgenomen.

Bij de kinderen met communicatieve handicaps blijkt de afname van de SON-R 6-40 goed mogelijk. De samenwerking en het begrip van de instructies wordt bij vrijwel alle kinderen door de test-leider als goed beoordeeld. Voor de motivatie en concentratie is dit ongeveer 90%. Ten opzichte van kinderen in de leeftijd van de SON-R 2½-7 blijkt de testafname veel gemakkelijker te verlopen, daar had ongeveer 40% van de kinderen problemen met de concentratie.

Voor personen met communicatieve beperkingen is een niet-verbale test noodzakelijk om tot een valide beoordeling van het intelligentieniveau te komen aangezien achterstand in de verbale ontwikkeling geheel of gedeeltelijk een gevolg kan zijn van de handicap en geen relatie hoeft te hebben met andere aspecten van de cognitieve ontwikkeling. In deze groep hangt het SON-IQ duidelijk samen met het de totaalscore op de WISC-III en de NIO, maar de gemiddelde scores op de SON-R 6-40 zijn aanmerkelijk hoger. Vooral op de verbale delen van de WISC-III en de NIO scoren de kinderen met communicatieve beperkingen laag. De SON-R 6-40 heeft correlaties van rond de .40 met indicatoren van schoolprestaties en doet het daarin beter dan de WISC-III en de NIO.

### **Zeer moeilijk lerende kinderen**

Het onderzoek bij kinderen met een ontwikkelingsachterstand is uitgevoerd bij kinderen op een school van het speciaal basisonderwijs en een school voor voortgezet speciaal onderwijs. Hier zijn relatief veel kinderen onderzocht die eerder niet of moeilijk testbaar bleken. Bij deze groep kinderen is vaak sprake van een veelvoud van sociale -, emotionele - en gedragsproblemen, en achterstanden in de cognitieve, verbale en motorische ontwikkeling.

Het gemiddelde SON-IQ van de zeer moeilijk lerende kinderen ligt rond de 70. Op alle onderdelen van de test is er sprake van een grote achterstand. Veertig procent heeft een SON-IQ lager dan zestig en scores boven de honderd komen niet voor. De prestaties op de test komen redelijk overeen met het schooloordeel over de intelligentie en de taalvaardigheid (circa .40) en nog meer met het oordeel over concentratie (.52). Deze correlaties zijn vergelijkbaar met het PIQ maar zijn lager dan voor het VIQ van de WISC-III. De correlatie tussen de SON-R 6-40 en de WISC-III is bijzonder hoog (voor het TIQ .86) maar de gemiddelde score van het TIQ is bijna tien punten lager dan het SON-IQ. Indien rekening wordt gehouden met de lagere ondergrens bij de WISC-III dan is dit verschil nog zeven IQ-punten.

De kinderen in deze groep zijn moeilijk testbaar. In vergelijking tot de kinderen met communicatieve handicaps is er naar het oordeel van de testleider vaker sprake van een matige of wisselende motivatie, concentratie, samenwerking of begrip van de instructies. In ongeveer de helft van de gevallen hebben deze kinderen problemen om de instructies te begrijpen en bij een derde is er sprake van concentratieproblemen. Toch bleek de afname van de test wel goed te doen terwijl bij een deel van deze kinderen de afname van een algemene intelligentietest eerder niet was gelukt.

### **Moeilijk testbare kinderen**

In een ideale testsituatie zijn kinderen goed gemotiveerd om de test te doen, begrijpen ze de instructies en zijn ze tot het eind geconcentreerd met de opgaven bezig. De praktijk kan echter anders zijn. Bij kinderen, en ook bij volwassenen, kan niet verwacht worden dat de motivatie en de concentratie bij voorbaat aanwezig zullen zijn. De testsituatie, het materiaal en de interactie met de testleider moeten zodanig zijn dat de cliënt geïnteresseerd raakt, en het verloop van de afname dient zo gestructureerd te zijn, dat de belangstelling ook behouden blijft.

Verschillende kenmerken spelen een rol waardoor de SON-tests in het algemeen goed zijn af te nemen. Het niet-verbale karakter van de test, waarbij de cliënt kan, maar niet hoeft te spreken, is aantrekkelijk voor personen die verlegen zijn of die wantrouwend tegenover de testsituatie staan. De adaptieve procedure en de feedback steunen de cliënt in het maken van de opgaven. Dit vermindert het gevoel te falen, helpt om de bedoeling van de opgaven duidelijk te maken en leidt tot een meer natuurlijke interactie tussen testleider en de cliënt. De sterke inperking van het aantal items dat fout wordt gemaakt, en de continuering met gemakkelijkere opgaven na een beperkt aantal fout gemaakte items voorkomt dat de cliënt snel gedemotiveerd raakt. Verder is er bij de SON-R 6-40 sprake van gevarieerde testonderdelen. Bij een aantal van deze testonderdelen dient de cliënt ook actief te handelen om tot de goede oplossing te komen.

Deze aspecten maken de SON-R 6-40 aantrekkelijk voor gebruik bij personen waarbij wellicht geen sprake is van een specifieke communicatieve handicap, maar waarbij sociale -, emotionele - en gedragsproblemen de afname van een meer traditionele intelligentietest bemoeilijken.

### Verstandelijk gehandicapten

De afname en interpretatie van intelligentietests bij verstandelijk gehandicapten is moeilijk omdat er zo'n grote discrepantie bestaat tussen de leeftijd en het niveau van functioneren. Wanneer een test wordt afgenomen die aansluit bij de leeftijd, dan zullen de genormeerde scores vaak de laagste waarde hebben en is een verdere differentiatie naar niveau niet mogelijk. Echter, wanneer een test wordt gekozen die aansluit bij het niveau van de persoon, zoals bijvoorbeeld een test voor jonge kinderen, dan kan wel een referentieleeftijd worden berekend, maar geen IQ-score of andere voor leeftijd genormeerde standaardscores (Tellegen, 2002d). De afname van een test die aan het niveau van de persoon is aangepast, zal dan toch vaak de voorkeur hebben. De wijze van afname en het moeilijkheidsniveau van de opgaven sluiten dan goed aan bij de mogelijkheden van de persoon in kwestie, en de afname zal motiverender zijn dan de afname van een test van een (te) moeilijk niveau.

Door de brede leeftijdsrange van de SON-R 6-40 is het mogelijk om beide aspecten van de meting te combineren. Indien het niveau waarop de cliënt functioneert niet veel lager is dan van een zesjarige dan kan zowel het IQ als de referentieleeftijd worden bepaald. De ondergrens voor het berekenen van de referentieleeftijd is namelijk 5;6 jaar. Indien het niveau hier naar alle waarschijnlijkheid onder zit dan kan men overwegen de SON-R 2½-7 af te nemen.

Wanneer er geen sprake meer is van een sterke ontwikkeling, kunnen de referentieleeftijden van verstandelijk gehandicapten goed met elkaar vergeleken worden. Indien het onderzoek echter verstandelijk gehandicapte kinderen betreft die in ontwikkeling zijn, dan moet men de onderlinge vergelijking van de referentieleeftijden beperken tot personen van ongeveer dezelfde leeftijd.

Met de SON-R 2½-7 en de SON-R 5½-17 is op beperkte schaal onderzoek gedaan bij kinderen en volwassenen met een verstandelijke handicap (Wijnands, 1997). De referentieleeftijd van de SON-tests bleek .79 te correleren met de referentieleeftijd gebaseerd op diverse andere tests, waaronder verschillende versies van de Wechsler tests. De positieve punten van de SON-R 6-40 die hiervoor genoemd zijn met betrekking tot moeilijk testbare kinderen, blijken ook bij het testen van personen met een verstandelijke handicap van groot belang. Bij deze groep personen bestaat vaak een grote angst te falen, en de feedback en de adaptieve procedure, waardoor na twee fouten met meer gemakkelijke items verder wordt gegaan, dragen bij tot de motivatie en het plezier om de test te maken.

### Allochtone kinderen en volwassenen

Het testen van allochtone kinderen en volwassenen met traditionele intelligentietests kan leiden tot grote onderschatting van hun cognitieve mogelijkheden omdat er veelal geen rekening wordt gehouden met het feit dat gebrekkige kennis van de taal waarin de testafname plaatsvindt geenszins hoeft in te houden dat de verbale intelligentie van de persoon beperkt is. Lagere prestaties op het verbale deel kunnen, maar hoeven niet te duiden op een lagere intelligentie. Ook kan de prestatie op het performale deel van deze tests *biased* zijn omdat de instructies bij traditionele tests standaard op een verbale wijze gegeven worden. Uit vergelijkingen waarbij de SON-R 2½-7 en een aantal andere tests bij dezelfde groepen zijn afgenomen, blijkt dat de verschillen tussen autochtone en allochtone kinderen in de meeste gevallen bij de SON-R 2½-7 minder groot zijn. Dergelijke analyses zijn voor de SON-R 6-40 nog niet verricht.

Het is wel mogelijk om het gemiddeld IQ van de totale groep allochtonen bij de SON-R 6-40 te vergelijken met de gemiddelde score van de allochtonen in het Nederlandse deel van het normeringsonderzoek van de WISC-III (Wechsler, 2005a). Het gemiddelde SON-IQ is voor de allochtonen

96.3 (N=184) en bij de WISC-III is het TIQ voor deze groep gemiddeld 91.9 (N=103). Dit betekent dat de achterstand van allochtonen bij de WISC-III ruim twee keer zo groot is. Bovendien zijn bij de normering van de WISC-III, vanwege het verbale karakter van de test, allochtonen die minder dan zes jaar in Nederland woonden van de normering uitgesloten. Indien deze groep wel zou zijn getest dan zou het verschil tussen de WISC-III en de SON-R 6-40 naar alle waarschijnlijkheid nog een stuk groter zijn geweest. Het verschil tussen het SON-IQ en het TIQ van de WISC-III kan niet zonder meer verklaard worden uit het ontbreken van specifiek verbale onderdelen bij de SON-test. Het gemiddelde PIQ van de allochtonen is bij het normeringsonderzoek van de WISC-III namelijk 92.1 en ten opzichte van het populatiegemiddelde van 100 twee keer zo laag als het gemiddelde SON-IQ. De verbale wijze van instructies bij de performale onderdelen en de demotiverende afbreekregels bij de WISC-III kunnen er toe bijdragen dat zowel de verbale als de performale scores op deze test zo laag zijn, zelfs wanneer de afname beperkt blijft tot het performale deel en wanneer alleen allochtone kinderen worden getest die tenminste zes jaar in Nederland wonen. Ook in deze gevallen blijkt de WISC-III discriminerend te werken en het gebruik van deze test bij allochtonen dient te worden ontraden. Alhoewel psychologen zich in het algemeen wel degelijk bewust zijn van de beperkingen die het testen van allochtonen met zich meebrengt, wordt hier in de praktijk vaak nog onvoldoende rekening mee gehouden (Zumdick, 2011).

Bij de SON-R 6-40 is bij allochtonen eigenlijk alleen van achterstand sprake indien beide ouders in het buitenland zijn geboren. Bij deze personen is het opleidingsniveau van de ouders laag hetgeen voor een deel de lage scores verklaart. In het normeringsonderzoek van de SON-R 6-40 zijn geen aanwijzingen gevonden dat de inhoud van de plaatjes van Categorieën problemen geeft voor personen met een andere culturele achtergrond. Integendeel, de allochtonen waarvan beide ouders in het buitenland zijn geboren behalen op deze subtest gemiddeld de hoogste score.

Alle uitkomsten wijzen op een goede bruikbaarheid van de SON-R 6-40 bij allochtonen. Natuurlijk kan het belangrijk zijn ook informatie te hebben over de taalvaardigheid en van de kennis die een allochtoon heeft van de Nederlandse taal. Dit moet echter niet verward worden met intelligentie, en de taalvaardigheid dient de beoordeling van de intelligentie niet direct te beïnvloeden. Verbale intelligentie dient men te beoordelen aan de hand van tests die in de moedertaal zijn afgenomen (Carroll, 1993).

### Visuele handicaps

De SON-R 6-40 is sterk visueel georiënteerd. Alle subtests maken gebruik van plaatjesmateriaal. Indien er sprake is van een sterk verminderd gezichtsvermogen en dit niet met een hulpmiddel gecompenseerd wordt, moet het gebruik van de SON-tests worden afgeraden. Voor deze categorie personen zijn wel aangepaste tests beschikbaar (Dekker, 1987). Geringe beperkingen van het gezichtsvermogen zullen waarschijnlijk geen probleem zijn. De afbeeldingen in de test zijn groot en duidelijk en doen geen beroep op het discrimineren van kleine visuele verschillen.

### Motorische handicaps

Verschillende taken van de SON-R 6-40 doen een beroep op motorische vaardigheden en oog-hand-coördinatie. Bij de constructie van de test is er naar gestreefd de invloed hiervan op de beoordeling van de prestaties te minimaliseren. Bij het onderdeel Patronen worden ruime criteria voor het beoordelen van de tekeningen gehanteerd. Bij Mozaïeken wordt met kaders gewerkt om het motorisch onhandige cliënten gemakkelijker te maken de opgaven goed te doen. Verder zijn, voor zover daarmee gewerkt wordt, de tijdslimieten ruim en wordt er niet op snelheid gescoord. Voor personen met ernstigere motorische handicaps dient men er echter rekening mee te houden dat deze handicaps de prestaties nadelig kunnen beïnvloeden (van de Beek, 1995). In het deel Instructies van de



handleiding wordt ingegaan op enkele mogelijkheden om de afnameprocedure aan te passen aan motorische problemen van de cliënt.

### **Gebruik van de test in andere landen**

Door het niet-verbale karakter van de test, kan deze gemakkelijk in andere landen worden toegepast. Het normeringsonderzoek in Duitsland en China heeft uitgewezen dat het testmateriaal althans in deze landen goed bruikbaar is.

Wanneer de test gebruikt wordt in landen en culturen die in sociaal-economisch opzicht nog weinig ontwikkeld zijn en in dit opzicht sterk van Nederland afwijken, zal men moeten nagaan in hoeverre het testmateriaal en de testsituatie voldoende herkenbaar zijn en in hoeverre aanpassing is vereist. Indien de omstandigheden waarin kinderen opgroeien heel armoedig zijn en de omgeving weinig mogelijkheden tot ontwikkeling biedt dan staat het testmateriaal zo ver van de leefomstandigheden af dat men niet kan verwachten dat de prestaties op de test een goed beeld geven van de cognitieve capaciteiten (Tellegen & Laros, 2005b). Dit kan ook gelden indien men in Nederland personen test die uit een andere cultuur komen of in Nederland in sterk afwijkende omstandigheden zijn opgegroeid.

### **Leeftijden**

Bij de constructie van de SON-R 6-40 is uitgegaan van een leeftijdsbereik van de test van zes tot veertig jaar. De normen van de test zijn geconstrueerd voor de leeftijd van vijf jaar en negen maanden tot de leeftijd van veertig jaar. De SON-R 2½-7 is genormeerd tot acht jaar. Hieronder zullen we voor een aantal leeftijdsgroepen aangeven hoe de SON-R 6-40 kan worden gebruikt en wanneer de SON-R 2½-7 beter kan worden gebruikt.

**jonger dan 5;9 jaar** Voor deze leeftijd is alleen de SON-R 2½-7 geschikt aangezien er geen normen zijn voor de SON-R 6-40. Indien de normen berekend worden met het computerprogramma dan worden deze berekend voor de leeftijd van 5;9 jaar. Dit wordt in het programma en op het scoreformulier aangegeven.

**5;9 – 6;0 jaar** Bij voorkeur zal men voor deze leeftijd de SON-R 2½-7 gebruiken hoewel er met het computerprogramma voor de SON-R 6-40 wel genormeerde scores kunnen worden berekend.

**6;0 – 6;11 jaar** Zowel de SON-R 2½-7 als de SON-R 6-40 zijn voor deze leeftijd geschikt. De SON-R 2½-7 heeft geringe plafondefecten. Voor onderzoek van (hoog)begaafde kinderen kan men beter de SON-R 6-40 gebruiken. Voor kinderen met cognitieve achterstanden en met handicaps, en voor moeilijk testbare kinderen, kan de SON-R 2½-7 beter worden gebruikt.

**7;0 – 7;11 jaar** In het algemeen kan men voor deze leeftijd beter de SON-R 6-40 gebruiken. De betrouwbaarheid en de generaliseerbaarheid van de SON-R 2½-7 zijn wel goed, maar er zijn duidelijke plafondefecten. Voor kinderen op een beneden gemiddeld niveau hoeft dit echter geen probleem te zijn. De afname van de SON-R 2fh,-7 kan in deze leeftijdsgroep aantrekkelijk zijn voor kinderen met handicaps, kinderen met een lage cognitieve ontwikkeling en voor moeilijk testbare kinderen die kunnen profiteren van de hulp die bij de test geboden wordt en van het gemakkelijke niveau van de eerste items.

**8;0 – 40;0 jaar** Standaard zal men voor deze leeftijd de SON-R 6-40 gebruiken. Vanaf deze leeftijd zijn er voor de SON-R 2½-7 geen normen voor de standaardcores. De SON-R 6-40 is genormeerd tot de leeftijd van 40;0 jaar. Voor personen vanaf acht jaar kan de SON-R 2½-7 interessant zijn wanneer



het niveau zo laag is dat de afname van een test die aansluit bij de mogelijkheden van de persoon praktisch beter uitvoerbaar is, en waarbij een goede bepaling van de referentieleeftijd informatiever is dan een extreem lage IQ-score.

**ouder dan 40;0 jaar** Indien een volwassene van veertig jaar of ouder wordt getest dan worden de normen gebaseerd op de leeftijd van 40;0 jaar. Dit wordt in het programma en op het scoreformulier aangegeven. Tot de leeftijd van ongeveer 45 jaar zullen de normen nog niet veel veranderen maar naarmate men ouder wordt zullen de prestaties in de populatie afnemen en zijn de IQ-scores die op de leeftijd van 40 jaar worden gebaseerd te laag (Wichers, 2005). Vanaf ongeveer 45 jaar geeft de uitkomst dus niet meer goed weer hoe men in vergelijking tot leeftijdsgenoten presteert. De referentieleeftijd die een range heeft van 5;6 tot 18;0 jaar kan wel normaal worden berekend. Bij ouderen die op een laag niveau functioneren kan dit zinvolle informatie zijn.

## 11.5 DE INTERPRETATIE VAN DE SCORES

De belangrijkste functie van de afname van een intelligentietest zoals de SON-R 6-40 is het verschaffen van informatie over het niveau van de cognitieve ontwikkeling van een cliënt ten behoeve van diagnostiek, advies en hulpverlening en voor (tussentijdse) evaluatie van het effect van behandelingsprogramma's en interventies. Dit betekent dat de uitkomsten op de test grote gevolgen kunnen hebben voor de leefsituatie van de cliënt en voor zijn of haar ontwikkeling. Ook voor de ouders of partner kunnen de gevolgen ingrijpend zijn, en adviezen of besluiten die in het kader van de diagnostiek genomen worden kunnen grote financiële consequenties hebben.

In het algemeen zal de afname van de test niet geïsoleerd plaatsvinden, maar in het kader van observaties en gesprekken met een cliënt en mogelijk met zijn omgeving zoals partner, ouders, leerkrachten, huisartsen of de werkgever. Veelal wordt een intelligentietest in combinatie met andere cognitieve tests en persoonlijkheidstests afgenomen. De afname van de test kan hierbij plaatsvinden in het kader van een cyclus waarbij vraagstelling en het verzamelen van relevante informatie elkaar afwisselen (Kievit & Tak, 1996).

De SON-R 6-40 geeft op verschillende niveaus (subtests, totaalscore) en op verschillende wijzen (referentieleeftijd, deviatiescores, observaties) informatie over het functioneren. Hieronder wordt ingegaan op de waarde van deze informatie voor de diagnostiek, en de risico's die er zijn wanneer men de eenmalige uitkomst op een test interpreteert als 'het' intelligentieniveau.

### Type informatie

De doelstelling van de SON-R 6-40 is een indruk te geven van het algemene intelligentieniveau van de cliënt. De reden dit te doen aan de hand van heel verschillende subtests is niet om onderlinge verschillen in prestatie tussen de subtests te kunnen vaststellen. Diverse subtests worden gebruikt omdat de invloed van specifieke kenmerken van subtests op de totaalscore kleiner wordt wanneer de test uit meerdere subtests is opgebouwd. De nauwkeurigheid van de IQ-score wordt ook niet primair beoordeeld op grond van de betrouwbaarheid van de test, maar op grond van de generaliseerbaarheid. Bij de generaliseerbaarheid wordt alle variantie die specifiek is voor de subtests als niet relevant beschouwd. Het SON-IQ, met het 80%-waarschijnlijkheidsinterval dat op de generaliseerbaarheid is gebaseerd, dient bij de beoordeling van de testuitkomst de basis te zijn.

### Subtestscores

De verschillen tussen de scores op de subtests zijn het minst betrouwbaar en stabiel. Uit het hertestonderzoek blijkt dat ook de mate waarin de subtestscores van elkaar verschillen, geen stabiel ken-

merk is. Wanneer de onderlinge verschillen relatief groot zijn is de kans wel groter dat de rangorde van de verschillen grotendeels behouden blijft. Wanneer men verschillen tussen de subtest scores nader wil interpreteren, is het nodig eerst na te gaan of de onderlinge verschillen voldoende groot zijn. Het computerprogramma biedt hiervoor de mogelijkheid.

Hoewel het verstandig is geen vergaande conclusies te verbinden aan uitkomsten op subtestniveau, kan het de moeite waard zijn om verschillen tussen subtest scores te evalueren in samenhang met andere informatie die men over de cliënt heeft, of met indrukken die tijdens de afname van de test zijn opgedaan. Op grond hiervan kunnen specifieke ideeën worden ontwikkeld over sterke en zwakke kanten van de cliënt, die vervolgens verder onderzocht kunnen worden. Het gaat hierbij om exploratief gebruik van de resultaten op subtestniveau dat zinvol is wanneer de onderlinge verschillen voldoende groot zijn.

### Het SON-IQ

Het SON-IQ, de genormeerde en gestandaardiseerde totaalscore van de SON-R 6-40, is de best bruikbare, generaliseerbare, en meest betrouwbare en stabiele uitkomst op de test. Tezamen met het 80%-waarschijnlijkheidsinterval geeft het SON-IQ een goede indicatie van het intelligentieniveau van de cliënt. Wanneer men de testuitkomst globaal wil omschrijven, kan men gebruik maken van de categorieën zoals die in Tabel 11.2 in de eerste kolom worden vermeld. Deze indeling is neutraal en puur beschrijvend van aard: aangegeven wordt of de cliënt, hoog, laag of gemiddeld op de test heeft gepresteerd. Bij deze indeling wordt een karakteristiek gegeven van de prestatie maar niet van de persoon. De tweede indeling wordt in de Verenigde Staten wel bij de Wechsler tests gebruikt (Sattler, 1992). Hoewel de grenzen voor de vier hoogste categorieën één IQ-punt verschillen komt deze indeling verder overeen met de zeven categorieën die ook bij de eerste indeling worden gebruikt. Er bestaat echter een essentieel verschil: de omschrijving van de eerste vijf categorieën kan nog beschouwd worden als aanduiding van de prestatie maar bij de laagste twee categorieën wordt een diagnose gesteld. Dit geldt ook voor de categorieën in de indeling van Resing en Blok (2002) die door de Cotan wordt geadviseerd. Hier wordt de persoon gekarakteriseerd en gediagnosticeerd op grond van de test score. Hiermee wijkt deze indeling wezenlijk van de eerste indeling af.

Het gebruik van beide laatste indelingen moet sterk ontraden worden. Er wordt namelijk ten onrechte van uitgegaan dat scores op intelligentietests onderling inwisselbaar zijn (Van Toorn & Bon, 2011). Zowel meetfouten als essentiële verschillen in normpopulaties compliceren de onderlinge inwisselbaarheid van IQ-scores. Een voorbeeld hiervan is de Rakit (Bleichrodt, Drenth, Zaal & Resing, 1984) waar het speciaal onderwijs buiten de normpopulatie is gelaten. Deze omissie had tot gevolg dat 10% van de leerlingen in het basisonderwijs volgens de categorieën die de auteurs hanteren als

**Tabel 11.2** Indeling van IQ-scores en intelligentieniveaus

IQ	omschrijving (1)	pct.	IQ	omschrijving (2)	IQ	omschrijving (3)
>130	zeer hoog	2%	>129	very superior	>130	zeer begaafd
121-130	hoog	7%	120-129	superior	121-130	begaafd
111-120	boven gemiddeld	16%	110-119	high average	111-120	boven gemiddeld
90-110	gemiddeld	50%	90-109	average	90-110	gemiddeld
80-89	onder gemiddeld	16%	80-89	low average	80-89	beneden gemiddeld
70-79	laag	7%	70-79	borderline	70-79	laag begaafd
<70	zeer laag	2%	<70	mentally deficiënt	<70	zwakzinnig

(1) indeling bij de SON-tests (Tellegen & Laros, 1998); (2) indeling bij de Wechsler-schalen (Sattler, 1992); (3) door de Cotan aanbevolen indeling (Resing & Blok, 2002)

‘debiel’ of ‘zwakzinnig’ zou zijn gediagnosticeerd (Tellegen, 2002c). Maar ook wanneer de normen overeenkomen en tests .80 correleren is de kans ruim dertig procent dat de scores tien of meer punten van elkaar verschillen. Daarnaast speelt het Flynn-effect een belangrijke rol en kan het gebruik van leeftijdsintervallen bij de normering, zoals bij de Wechsler tests, tot grote verandering in het IQ leiden indien de test een dag eerder of later wordt afgenomen. Op grond hiervan moeten we concluderen dat we niet kunnen spreken van ‘het’ IQ van een persoon (Tellegen, 2004). Het IQ is puur de getalsmatige uitkomst op een bepaalde test op een bepaald moment. Het is de taak van de diagnostiek om de betekenis van deze uitkomst te interpreteren. Bij de beoordelingssystemen van de Wechsler tests en de Cotan is de score echter al de diagnose, waarbij IQ-scores op verschillende tests ten onrechte als inwisselbaar worden beschouwd. De doelstelling van de SON-tests is daarentegen dat het SON-IQ gebruikt wordt als hulpmiddel bij de diagnostiek om tot een zo goed mogelijke beoordeling en advisering te komen.

### *De referentieleeftijd*

De uitkomsten op de test kunnen op het niveau van de subtests en voor de totale testuitkomst ook worden weergegeven als referentieleeftijd. Hierbij geldt, evenals voor de standaardscores, dat de referentieleeftijd die gebaseerd is op de totale testuitkomst het meest betrouwbaar, stabiel en generaliseerbaar is, en daarom voor de beoordeling van de uitkomsten het best gebruikt kan worden.

Voor volwassenen boven de veertig jaar, is de referentieleeftijd de enige genormeerde informatie die beschikbaar is. Voor personen in de leeftijd van zes tot veertig jaar kan de referentieleeftijd gebruikt worden als aanvullende informatie bij het SON-IQ.

Over het nut van referentieleeftijden (ook wel testleeftijd of mentale leeftijd genoemd) wordt verschillend gedacht (voor kritiek op het gebruik van de verwante didactische leeftijd, zie Evers en Resing, 2007). De referentieleeftijd geeft een absoluut niveau van presteren weer. In tegenstelling tot de standaardscores, die de positie van de prestatie weergeven ten opzichte van leeftijdsgenoten, is de berekening van de referentieleeftijd niet van de leeftijd van de cliënt afhankelijk. De interpretatie van de referentieleeftijd dient echter wel te gebeuren in samenhang met de leeftijd bij afname van de test. Een referentieleeftijd van 8;3 jaar betekent voor een kind van 6;1 jaar en een kind van 11;2 jaar dat zij even goed op de test hebben gepresteerd, echter de psychologische betekenis van de testuitkomst is voor beide kinderen geheel verschillend.

Van de IQ-score kan verwacht worden dat deze in de loop van de ontwikkeling min of meer stabiel is. Dit geldt echter niet voor de referentieleeftijd; zolang de persoon zich ontwikkelt zal ook de referentieleeftijd toenemen. Aangezien de ontwikkeling op jongere leeftijd veel sneller verloopt dan op oudere leeftijd, zal voor een cliënt de discrepantie tussen de referentieleeftijd en de leeftijd bij afname van de test steeds groter worden. Op de leeftijd van 8;0 jaar is bij een IQ van 80 de discrepantie in maanden veel kleiner dan op de leeftijd van 12;0 jaar. Verder is, gegeven een vaste leeftijd, voor een IQ van 80 de discrepantie kleiner dan voor een IQ van 120. Dit maakt dat referentieleeftijden, of discrepanties tussen referentieleeftijd en leeftijd bij afname van de test, moeilijk met elkaar vergeleken kunnen worden en zich ook niet goed lenen voor statistische analyse.

Een ander nadeel bij het gebruik van de referentieleeftijd is dat hierbij niet, zoals bij de IQ-score, een waarschijnlijkheidsinterval wordt geboden dat aangeeft hoe (on)nauwkeurig de uitspraak over de referentieleeftijd is.

Ondanks deze beperkingen kan de referentieleeftijd zeker nuttig zijn. De referentieleeftijd geeft op een heel concrete wijze weer hoe de cliënt functioneert op de test en bij rapportage van de uitkomsten naar bijvoorbeeld ouders kan deze informatie verhelderend zijn. Verder geeft de referentieleeftijd informatie over het niveau van opgaven dat de cliënt begrijpt en dit kan een aanwijzing geven op welk niveau lesmateriaal of trainingen gegeven kunnen worden. Hierbij kan men natuurlijk niet alleen op de referentieleeftijd afgaan. Wanneer een twaalfjarig kind een referentieleeftijd van

6;5 jaar heeft, is dit kind in een geheel andere situatie, en heeft het heel andere leermogelijkheden, dan een zesjarige met een referentieleeftijd van 6;5 jaar. Overigens is een zesjarige met een IQ van 80 natuurlijk ook niet 'gelijk' aan een twaalfjarige met een IQ van 80. De referentieleeftijd kan men het best omschrijven als: *de prestatie van de cliënt op de test komt overeen met de gemiddelde prestatie van personen van .. jaar*. Dit is beter dan wanneer men zegt: *de cliënt functioneert op het niveau van een .. jarige*. De laatste formulering suggereert ten onrechte dat met de test het hele cognitieve of mentale niveau van de cliënt wordt beschreven.

### **Oordeel testleider**

In het onderzoek met de test zijn de deelnemers door de testleider na afloop beoordeeld op de aspecten motivatie, concentratie, samenwerking en begrip van de instructie. Bij personen met ontwikkelingsachterstanden en stoornissen, en ook bij jongere kinderen, komt het regelmatig voor dat er met betrekking tot deze aspecten problemen zijn, en vaak zijn dit ook kinderen die minder goed of slecht op de test presteren. Wanneer deelnemers wel in staat waren de test te doen, is een negatief oordeel geen reden geweest hen bij de beschrijving en analyse van de uitkomsten buiten beschouwing te laten. Voor de diagnostiek kan deze informatie echter wel van betekenis zijn. Wanneer de indruk bestaat dat de cliënt slecht gemotiveerd en geconcentreerd was, zal men zich moeten afvragen of de testuitkomst wel een valide indicatie van de intelligentie geeft. Hierbij is het belangrijk of de motivatie- en concentratieproblemen zich 'toevallig' bij deze testafname voordeden, of dat dit problemen zijn die kenmerkend zijn voor de cliënt en die zich in tal van situaties voordoen. De vraag kan dan gesteld worden of de aanpak van deze problemen op den duur tot betere test- en leerprestaties zullen leiden.

Wanneer iemand op de test slecht presteert en het oordeel met betrekking tot de verschillende aspecten is positief, dan kan men er meer vertrouwen in hebben dat de cliënt wel heeft laten zien waartoe hij of zij in staat is, en dat de lage score niet het gevolg is van moeilijke testbaarheid van de cliënt.

De vier beoordelingscategorieën die in het onderzoek gebruikt zijn, en die ook op het scoreformulier staan afgedrukt, kunnen als uitgangspunt dienen bij de observatie. Omdat cliënten bij de SON-R 6-40 zelf actief bezig zijn, is er veel gelegenheid tot observatie en we verwachten dat gebruikers van de test daar gebruik van kunnen maken.

### **Generaliseren van de testuitkomst**

Het SON-IQ geeft weer hoe goed de cliënt op de test heeft gepresteerd. Op grond van de eerste indeling in Tabel 11.2 kan men deze prestatie in de categorieën van 'zeer hoog' tot 'zeer laag' indelen. Wanneer men met de beperkingen van de test rekening wilt houden, en op grond van de prestatie een algemeen geldende uitspraak wil doen over het intelligentieniveau is dit lastiger.

### **Generaliseren over subtests**

Het 80%-waarschijnlijkheidsinterval dat steeds bij de IQ-scores wordt gegeven, houdt met twee beperkingen rekening: namelijk met de onbetrouwbaarheid van de test en met het feit dat een deel van de betrouwbare variantie specifiek is voor de verschillende subtests. Het interval geeft aan waar naar verwachting de IQ-score zou liggen indien een zeer groot aantal vergelijkbare subtests zou zijn afgenomen. Deze score zou dan vrijwel perfect betrouwbaar zijn en de invloed van specifieke kenmerken van de subtests zou dan te verwaarlozen zijn. De totale breedte van dit interval is 14 à 15 punten. De meeste beschrijvende categorieën van de IQ-score in Tabel 11.2 zijn 10 punten breed. Dit betekent dat het 80%-interval van de IQ-score of twee categorieën omvat, of één categorie en nog een deel van beide aangrenzende categorieën. Wanneer de uitkomst op de test als intelligentiescore beschreven wordt, kan het best het 80%-interval vermeld worden, of, wanneer men van catego-

rieën zoals in Tabel 11.2 gebruik wil maken, een omschrijving van de verschillende categorieën die het interval bestrijkt.

### *Generaliseren over tijd*

Het 80%-interval van de IQ-score houdt geen rekening met de stabiliteit van de test over een periode van enkele maanden. Voor de beoordeling van de intelligentie is dit echter wel van belang. Gemiddeld is de betrouwbaarheid .953 en de generaliseerbaarheid .840. De stabiliteit ligt met .917 vrij dicht bij de betrouwbaarheid. Dit betekent dat het 80%-waarschijnlijkheidsinterval niet veel breder wordt als rekening wordt gehouden met wisselvalligheid in prestatie in de loop van enkele maanden. Een globale berekening geeft als uitkomst dat het interval aan beide zijden één punt breder wordt. Indien men dus bijvoorbeeld een 80%-waarschijnlijkheidsinterval heeft van 107-121 en bij de interpretatie rekening wilt houden met instabiliteit van de prestaties, dan kan men voor de inschatting van het intelligentieniveau beter uitgaan van een interval dat loopt van 106-122. Dit betekent dat deze persoon volgens de Cotan-classificatie zowel in de gemiddelde als in de begaafde categorie thuis zou kunnen horen. Een enkele testafname geeft onvoldoende informatie om daar een betrouwbare uitspraak over te kunnen doen.

### *Afname meerdere tests*

Wanneer de test gebruikt wordt om belangrijke beslissingen te nemen die grote gevolgen kunnen hebben voor de cliënt, kan het belangrijk zijn om bij de beoordeling niet te volstaan met de afname van één intelligentietest. Het risico is dan te groot dat door een combinatie van onbetrouwbaarheid, fluctuaties in de prestatie en specifieke kenmerken van de test, een vertekend beeld van de intelligentie ontstaat.

In paragraaf 11.2 is op grond van het onderzoek naar de congruente validiteit van de SON-R 6-40, de variantie van het SON-IQ als volgt beschreven (zie Figuur 11.2):

- meetfoutvariantie (5%),
- instabiele variantie (3%),
- testspecifieke variantie (12%),
- valide generaliseerbare variantie (80%).

De proportie valide generaliseerbare variantie is gebaseerd op correlaties van ongeveer .80 met andere intelligentietests. Wanneer we aannemen dat andere intelligentietests beschikbaar zijn waarvan de variantie op dezelfde wijze is samengesteld, en die met de SON-R 6-40 een correlatie hebben van .80, dan kan berekend worden hoe de variantie is samengesteld van de gemiddelde score wanneer twee tests worden afgenomen. Hierbij wordt ervan uitgegaan dat de periode tussen de afnames enkele weken tot enkele maanden kan bedragen. Bij de afname van twee tests neemt het aandeel van de ongewenste meetfoutvariantie en de instabiele variantie af tot 4% en wordt dus gehalveerd. De proportie valide generaliseerbare variantie neemt echter toe tot .89 waardoor het waarschijnlijkheidsinterval kleiner wordt. Een IQ-score die op meerdere test afnames is gebaseerd, kan dus met veel meer vertrouwen geïnterpreteerd worden als niveau van intelligentie.

Het afnemen van een andere test, in combinatie met de SON-R 6-40, is juist zo aantrekkelijk omdat dan tegelijkertijd de invloed van verschillende niet-gewenste bronnen van variantie vermindert wordt. Hierbij dient de alternatieve test natuurlijk wel geschikt te zijn voor de doelgroep. Voor diverse groepen cliënten waarbij men de SON gebruikt, zullen als alternatief alleen niet-verbale tests in aanmerking komen, of het performale deel van meer algemene tests. Diversiteit in materiaal en methode van testen is, voor zover mogelijk, gewenst. Ook is het aan te bevelen de tests door verschillende testleiders af te laten nemen. Omdat de SON-R 6-40 in methode van afnemen, en in

materiaal, zo duidelijk anders is dan veel andere tests, is de SON-R 6-40 zeer geschikt om als extra test te worden afgenomen bij cliënten waarbij een traditionele verbale intelligentietest is afgenomen.

### *Andere informatie*

Of het noodzakelijk is om naast de SON-R 6-40 ook een andere test af te nemen, is primair afhankelijk van de consequenties van een verkeerde beoordeling van de cliënt. Indien deze consequenties minder groot zijn, en wanneer een herziening van het oordeel gemakkelijk kan plaatsvinden, is een relatief grote onzekerheidsmarge acceptabel. Het risico dat een verkeerde beoordeling gegeven wordt zal ook kleiner zijn wanneer de uitkomst op de test in samenhang kan worden geïnterpreteerd met informatie van anderen die bij de cliënt betrokken zijn. De observaties van de testleider kunnen ook een indicatie geven of afname van een extra test gewenst is.

### **Wijze van afnemen**

Een voorwaarde voor de validiteit van de testuitkomst is dat de test op een goede manier en volgens de instructies wordt afgenomen. Hierbij is het van belang dat de testleider ervaren is in het afnemen van tests, en dat hij of zij ook ervaring heeft, indien daarvan sprake is, in de omgang met personen met specifieke problemen en handicaps. De afname van de test hoeft niet noodzakelijkerwijs door psychologen te gebeuren, de interpretatie en de rapportage van de uitkomsten blijft echter voorbehouden aan daartoe bevoegde deskundigen.

Een belangrijk aspect van de SON-R 6-40 is het cliëntvriendelijke karakter en de interactie tussen testleider en de persoon die wordt getest. Dit maakt de test zowel voor de cliënt als de testleider aantrekkelijk om te doen. Het betekent echter ook dat de testleider sterk bij de afname betrokken is en dat daardoor de risico's voor testleidereffecten groter zijn. Aanwijzingen voor systematische testleidereffecten zijn in ons onderzoek beperkt gebleven tot enkele IQ-punten. Om de risico's op het optreden van dergelijke effecten te verkleinen, is het aan te bevelen om naast het goed volgen van de instructies, ook eens aanwezig te zijn bij de afname van de test door iemand anders, en een ander eens mee te laten kijken wanneer men zelf de test afneemt. Indien dat mogelijk is, is het ook zinvol de uitkomsten die verschillende testleiders hebben verkregen, af en toe met elkaar te vergelijken.

## **11.6 CONCLUSIES**

Uit het onderzoek komt naar voren dat de SON-R 6-40 een betrouwbare en valide intelligentietest is die goed gebruikt kan worden bij kinderen en volwassenen met problemen en handicaps op het gebied van de taalontwikkeling en de communicatie, bij kinderen met een anderstalige of tweetalige achtergrond, bij kinderen met een ontwikkelingsachterstand en ontwikkelingsstoornissen, en bij zwakbegaafde kinderen en volwassenen. Wanneer bij deze groepen meer traditionele, verbale intelligentietests worden gebruikt kan het oordeel over de intelligentie te laag uitvallen door de beperkte taalvaardigheid van de cliënt of door de minder motiverende wijze van afnemen bij deze tests.

Kinderen en volwassenen met beperkte cognitieve capaciteiten of met problemen op het gebied van taal en communicatie zijn vaak moeilijk testbaar. In aanvulling op het niet-verbale karakter van de SON-R 6-40, waarbij de cliënt kan, maar niet hoeft te spreken, zorgt ook de adaptieve procedure, de aantrekkelijkheid van het materiaal, en de wijze waarop de cliënt zelf actief bezig kan zijn, voor een prettige testsituatie. In het algemeen komt naar voren dat de cliënten bij de SON-R 6-40 gemotiveerd en geconcentreerd zijn en dat zij de instructies goed begrijpen.

De scores op de test hebben een hoge samenhang met verschillende oordelen over intelligentie. Kinderen met een ontwikkelingsachterstand en leerproblemen presteren naar verwachting laag, maar veel beter dan op traditionele tests. De samenhang met andere intelligentietests is hoog. Door inhoudelijke verschillen tussen tests en door wisselvalligheid in de prestaties, zijn de correlaties echter lager dan op grond van de hoge betrouwbaarheid van de SON-R 6-40 theoretisch mogelijk is. De score op de test geeft een indicatie van de intelligentie van de cliënt; de score is niet 'het' intelligentieniveau. Bij ingrijpende beslissingen is het gewenst de diagnostiek te baseren op de afname van meerdere intelligentietests.

Primair voor de interpretatie van de testuitkomst is de IQ-score, waarvoor ook de referentieleeftijd kan worden bepaald. Door de normen voor de testcores op de exacte leeftijd van de cliënt te baseren, worden systematische vertekeningen bij de presentatie van de uitkomsten vermeden en door het vermelden van een waarschijnlijkheidsinterval kan de gebruiker rekening houden met de marge van onzekerheid van de uitkomsten.

Een aanzienlijk deel van de personen die voor psychologisch onderzoek in aanmerking komen, heeft problemen op het gebied van taal en communicatie. Daarnaast is een belangrijk deel van de bevolking niet Nederlandstalig opgegroeid, hetgeen een valide intelligentiebepaling met een traditionele verbale intelligentietest uitsluit. Gezien de grote diversiteit van de problemen en de handicaps in verband waarmee kinderen en volwassenen psychologisch worden onderzocht, is het van groot belang dat voor al deze groepen een goed geconstrueerde, genormeerde en gevalideerde intelligentietest beschikbaar is. Aan deze criteria voldoet de SON-R 6-40.