

De SON-R 6-40: stand van zaken

Peter Tellegen & Jacob A. Laros

Rijksuniversiteit Groningen & University of Brasilia

vierde versie, 26 mei 2011

Op 1 september 2011 zal de nieuwe versie van de SON-test voor de leeftijdsgroep van 6 tot 40 jaar, de SON-R 6-40, verschijnen. Omdat de SON-R 5½-17 vanwege veroudering van de normen vanaf het schooljaar 2011/2012 niet meer gebruikt mag worden in het kader van indicatiestelling, en omdat de beoordeling van de SON-R 6-40 door de COTAN pas kan plaatsvinden als de test is gepubliceerd, dreigde er een gat te vallen in de mogelijkheid de SON te gebruiken. Hierover is overleg geweest met de COTAN. Gezien de grote overeenkomst van de SON-R 6-40 met de verkorte vorm van de SON-R 5½-17 die in 2004 op alle aspecten als “goed” is beoordeeld, is het naar de mening van de COTAN voor verantwoord gebruik van de test voldoende als de nieuwe normen zijn goedgekeurd. Met de coördinator testbeoordeling van de COTAN is de afspraak gemaakt dat het hoofdstuk “Normering” van de handleiding van de SON-R 6-40 tussentijds beoordeeld wordt. Na de publicatie van de handleiding kan dan een volledige beoordeling van alle aspecten volgen. Zowel de REC als de RVC-VO hebben laten weten dat bij voldoende beoordeling van de normen, de SON-R 6-40 komend schooljaar in het kader van indicatiestelling gebruikt kan worden. Inmiddels zijn de normen door de COTAN als “goed” beoordeeld zodat er geen belemmeringen zijn voor het gebruik van de test.

Ten behoeve van het normeringsonderzoek is het testmateriaal al in definitieve vorm geproduceerd. Met de uitgever is de afspraak gemaakt om de handleiding in drie gedeelten uit te brengen (Verantwoording, Instructies, Normen) waardoor de productie sneller kan verlopen. Hierdoor zal de test 1 september 2011 kunnen verschijnen.

Aan de hand van de beoordelingspunten van de COTAN wordt hieronder een overzicht gegeven van de stand van zaken met betrekking tot de SON-R 6-40 zodat men een indruk kan krijgen van de kenmerken en kwaliteit van de opvolger van de SON-R 5½-17.

1. *Uitgangspunten van de testconstructie*

De doelstelling van de SON-R 6-40 is ongewijzigd gebleven: het geven van een betrouwbare en valide indruk van de algemene intelligentie zonder hierbij van het gebruik van taal afhankelijk te zijn. Bij het onderzoek naar de verkorte versie van de SON-R 5½-17 is gebleken dat door een goede selectie van vier subtests de validiteit, in termen van de relatie met onderwijsprestaties, nauwelijks verminderde. Bij de toenmalige verkorte vorm is gekozen voor de subtests Categorieën, Analogieën, Situaties en Mozaïeken. Bij de SON-R 6-40 is Patronen in de plaats gekomen van Situaties. Deze combinatie laat hogere correlaties zien met rapportcijfers en met het onderwijsniveau bij het voortgezet onderwijs dan de eerdere selectie. Deze correlaties zijn zelfs nog iets hoger dan bij afname van de volledige SON-R 5½-17.

Een tweede belangrijke doelstelling bij de herziening was het minder cultuurgebonden maken van de test. Met name voor de subtests met plaatjesmateriaal zou hiervoor uitgebreid crosscultureel onderzoek nodig zijn. Door Patronen te gebruiken in plaats van Situaties kon dit onderzoek beperkt blijven tot Categorieën. De SON-R 6-40 bestaat nu uit twee ruimtelijke handelingstests (Mozaïeken en Patronen) en twee redeneertests (Categorieën en Analogieën). Bij de ruimtelijke tests speelt logisch redeneren echter ook een grote rol. Anderzijds heeft de redeneertest Analogieën een sterk ruimtelijk karakter.

Om de moeilijkheidsopbouw van de subtests te verbeteren, en de test beter geschikt te maken voor personen met beperkte capaciteiten, is het aantal items per subtests uitgebreid. Hierdoor kan ook het verlies aan betrouwbaarheid vanwege de vermindering van het aantal

subtests beperkt worden. Bij de SON-R 6-40 bestaan Categorieën en Analogieën ieder uit drie parallelle reeksen van 12 items. Mozaïeken en Patronen bestaan uit twee reeksen van 13 items. Het totaal aantal items is nu 124, bij de SON-R 5½-17 was dit aantal bij deze subtests 98. Door de adaptieve procedure blijft de toename in afdameduur van de test beperkt. Deze is inclusief korte pauzes die soms tussen de subtests worden ingelast gemiddeld iets minder dan een uur.

Omdat er veel behoefte is aan een niet-verbale test voor volwassenen is het leeftijdsbereik bij de normering uitgebreid naar 40 jaar. Aan het karakter van de test is hierbij niets veranderd, wel zijn meer moeilijke items toegevoegd. Aan de onderzijde is het leeftijdsbereik beperkt tot 6 jaar. Voor 5-jarigen is de SON-R 2½-7 beter geschikt.

De meest ingrijpende veranderingen hebben plaatsgevonden met betrekking tot de subtest Categorieën. Omdat het plaatjesmateriaal was verouderd en de inhoud voor een deel cultuurgebonden (bijvoorbeeld de afbeelding van een slee en van een washand) zijn alle tekeningen nieuw gemaakt en is in diverse landen in Europa, Azië, Afrika en Zuid-Amerika onderzoek gedaan om dit testonderdeel beter cultuuronafhankelijk te maken.

Voor alle subtests is een belangrijke verandering in de wijze van afname dat nu bij de adaptieve procedure, bij de overgang naar een volgende reeks, één item verder wordt teruggegaan. Hierdoor is er meer zekerheid dat de overgeslagen items ook daadwerkelijk goed zouden zijn gemaakt indien afgenomen.

2. *Kwaliteit van het testmateriaal*

Het testmateriaal is door Hogrefe Verlag op dezelfde wijze, en met dezelfde kwalitatieve normen, geproduceerd als de SON-R 5½-17 en de SON-R 2½-7. Het materiaal zit in een rode kunststof koffer waarbij het verbruiksmateriaal en de scoreformulieren in het deksel worden opborgen. Dit maakt dat alle onderdelen gemakkelijk toegankelijk zijn. In vergelijking tot de SON-R 5½-17 is de test lichter, compacter en gemakkelijker hanteerbaar. De normen voor de SON-R 6-40 worden toegevoegd aan versie 5.1 van het SON computerprogramma. De CD zit standaard in de testkoffer.

Het scoreformulier is vergelijkbaar met het formulier van de SON-R 5½-17. De goede antwoordalternatieven zijn op het formulier aangegeven, evenals specifieke punten van de instructie en de regels voor de adaptieve procedure.

3. *Kwaliteit van de handleiding*

De handleiding wordt in drie delen uitgegeven: Verantwoording, Instructies en Normen. Het deel Instructies omvat de algemene instructies en de instructies per subtest. De verbale en niet-verbale instructies zijn naast elkaar afgedrukt.

De normtabellen zijn zo gedifferentieerd dat het verschil in IQ bij de overgang van de ene naar de andere leeftijdsgroep niet meer dan twee IQ-punten bedraagt. Voor het gebruiksgemak worden de normen voor de subtests en de totaalscore steeds naast elkaar afgedrukt. Overigens wordt sterk aangeraden om het computerprogramma te gebruiken. Dit biedt grotere nauwkeurigheid en extra mogelijkheden, zoals het berekenen van de referentieleeftijd en de mogelijkheid om een genormeerde totaalscore te berekenen voor elke combinatie van subtests. De genormeerde subtestscores worden evenals bij de WISC, WAIS en SON-R 2½-7 weergegeven met gemiddelde 10 en standaarddeviatie 3. Het SON-IQ heeft een gemiddelde van 100 en standaarddeviatie 15. De range is van 55-145. Naast dit standaard IQ wordt niet meer het specifiek en het gegeneraliseerd IQ berekend. Hoewel dit in psychometrisch opzicht voordelen biedt, bleek het in de praktijk verwarrend. Het onzekerheidsinterval dat bij de IQ-score wordt weergegeven, is evenals bij de SON-R 2½-7 gebaseerd op de generaliseerbaarheid en niet op de betrouwbaarheid.

Het deel Verantwoording geeft een overzicht van het vooronderzoek en de samenstelling van de test, het normeringsonderzoek en de betrouwbaarheid, en van het uitgebreide onderzoek naar de validiteit. Dit omvat onderzoek in Nederland en Duitsland en ook onderzoek dat in Brazilië, China en Ethiopië is uitgevoerd.

4. Normen

De normen van de SON-R 6-40 zijn gebaseerd op een steekproef van 1933 personen in de leeftijd van 6 tot 40 jaar waarvan 938 uit Nederland en 995 uit Duitsland. De steekproef is gestratificeerd naar leeftijd, regionale en gemeentelijke kenmerken en naar onderwijsniveau. Hier wordt eerst de uitvoering van het Nederlandse deel van het normeringsonderzoek besproken, daarna wordt ingegaan op de overeenkomst met de Duitse uitkomsten en de methode die is gevolgd om tot de genormeerde scores te komen.

Tabel 1

Globale opzet van de normgroep

Deel 1	Deel 2	Deel 3	Deel 4
Basisonderwijs	Voortgezet onderwijs	Volwassenen	Speciale groepen
-----	-----	-----	-----
Leeftijdsgroepen	Leeftijdsgroepen	Leeftijdsgroepen	verdeeld over alle leeftijden circa 5%
-----	-----	-----	
6 jaar (N=60)	12 jaar (N=60)	18 jaar (N=60)	
7 jaar (N=60)	13 jaar (N=60)	20 jaar (N=60)	
8 jaar (N=60)	14 jaar (N=60)	22 jaar (N=60)	
9 jaar (N=60)	15 jaar (N=60)	27 jaar (N=60)	
10 jaar (N=60)	16 jaar (N=60)	32 jaar (N=60)	
11 jaar (N=60)		37 jaar (N=60)	
-----	-----	-----	
(N=360)	(N=300)	(N=360)	

Bij de opzet van het Nederlandse onderzoek is uitgegaan van 17 leeftijdsgroepen van 60 personen met in iedere groep een gelijk aantal mannen en vrouwen, in totaal 1020 personen (zie Tabel 1). Van 6 t/m 16 jaar zijn het jaarlijkse groepen, daarna zijn de afstanden tussen de leeftijden groter omdat verwacht kan worden dat de veranderingen in prestatie bij volwassenen minder groot zijn. Bij de opzet van het onderzoek is onderscheid gemaakt tussen basisscholen (vnl. leeftijd 6-11 jaar), voortgezet onderwijs inclusief MBO (vnl. leeftijd 12-16 jaar) en volwassenen (de leeftijdsgroepen van 18 tot 40 jaar). Daarnaast zijn voor alle leeftijden proportioneel personen uit specifieke groepen toegevoegd zoals leerlingen van het speciaal (basis)onderwijs, praktijkonderwijs en volwassenen met een licht verstandelijke handicap. Naar rato van bevolking zijn in West-, Noord/Oost- en Zuid-Nederland tien zogenoemde onderwijsgebieden geselecteerd. De selectie van scholen voor voortgezet onderwijs vond binnen deze gebieden plaats. In de tien gebieden zijn 22 gemeenten geselecteerd waar het onderzoek op de basisscholen en bij de volwassenen is uitgevoerd. Deze gemeenten zijn representatief met betrekking tot gemeentegrootte, urbanisatiegraad, inkomens- en opleidingsniveau van de bevolking en met betrekking tot het percentage allochtonen. In Tabel 2 wordt een overzicht gegeven van de uiteindelijke normgroep met betrekking tot landsdeel, gemeentegrootte en stedelijkheid. De specifieke groepen zijn hier buiten gelaten. De landelijke percentages staan tussen haakjes.

Van de dertig geselecteerde basisscholen waren er 15 direct bereid mee te werken. Voor de weigeraars zijn 11 vergelijkbare vervangende scholen gevonden. De 26 scholen die hebben meegewerkt komen met betrekking tot het percentage leerlingen uit culturele minderheden en anderstaligen (15,4%) goed overeen met het landelijke percentage (13,7%). Binnen de school zijn leerlingen op leeftijd geselecteerd. Ongeacht groep of sekse dienden zij zo dicht mogelijk bij de gewenste leeftijd van X jaar en 6 maanden te zitten. Weigering van ouders of leerlingen kwam binnen het basisonderwijs nauwelijks voor.

Bij het voortgezet onderwijs was weigering van scholen een groot probleem. Van de 30 aangeschreven scholen waren slechts 6 bereid mee te werken. Na herhaalde aanschrijving van nieuwe scholen en met behulp van contacten van testleiders met scholen, kon dit deel van het onderzoek worden uitgevoerd bij 17 VO-scholen en 2 MBO-opleidingen. Door weigering of door praktische omstandigheden was het niet altijd mogelijk de geplande leerlingen te testen. In dat geval zijn vervangers van hetzelfde niveau en klas gezocht. Op grond van vergelijking van het onderwijsniveau met landelijke gegevens van het CBS zijn in de laatste fase van het onderzoek nog enkele leerlingen met specifieke kenmerken toegevoegd dan wel weggelaten om de steekproef beter representatief te maken.

Tabel 2

Samenstelling van de Nederlandse normgroep naar landsdeel, gemeentegrootte en stedelijkheid (exclusief bijzondere groepen; N=892)

landsdeel	gemeentegrootte		stedelijkheid			
	%	%	%	%		
Noord/Oost	33.6	(31.6)	< 50.000 inw.	46.6 (50.0)	niet/weinig	31.1 (33.9)
West	46.9	(46.8)	50-150.000 inw.	26.0 (28.5)	matig/sterk	48.8 (48.4)
Zuid	19.5	(21.6)	> 150.000 inw.	27.4 (21.5)	zeer sterk	20.1 (17.7)

Bij de volwassenen is in de 22 gemeenten uit het bevolkingsregister een steekproef getrokken van personen die in iedere leeftijdsgroep zo dicht mogelijk de gewenste leeftijd benaderden. Ondanks het feit dat men een vergoeding van € 50,- kreeg en de afname thuis kon plaatsvinden, reageerde slechts een kwart van de aangeschrevenen positief waardoor herhaalde aanschrijfrondes nodig zijn geweest en minder personen zijn getest dan gepland. Uit de gegevens die de volwassen deelnemers verstrekten, bleek dat de selectie had geleid tot een sterke oververtegenwoordiging van personen met een gemiddeld en hoog opleidingsniveau. Op grond hiervan zijn uiteindelijk 104 volwassenen uit de normgroep verwijderd.

Tabel 3

Onderwijsniveau naar leeftijd voor de normgroep en (tussen haakjes) voor de Nederlandse populatie

Onderwijstype	6-9 jaar		10-11 jaar		12-14 jaar		15-16 jaar	
	%	%	%	%	%	%	%	
	N=190		N=193		N=184		N=129	
Basisonderwijs	96.3	(96.5)	92.2	(92.8)	10.3	(14.1)		
S(B)O en PRO	3.7	(3.6)	6.8	(6.8)	6.6	(6.1)	4.7	(5.7)
Voortgezet onderwijs								
Vmbo bbl/kbl					25.5	(15.0)	22.5	(22.8)
Vmbo (g)tl					16.8	(12.5)	19.4	(18.5)
Havo					10.3	(8.0)	18.6	(20.8)
Vwo			1.0	(0.4)	12.5	(13.4)	21.7	(19.6)
Gemengd					17.9	(30.8)	1.6	(1.7)
MBO							11.6	(11.0)
LWOO (binnen VMBO)					11.5	(11.7)	13.8	(14.3)

In Tabel 3 is de samenstelling van de normgroep van 6-16 jaar naar onderwijsniveau weergegeven. Tussen haakjes zijn de populatiewaarden vermeld. Dat er relatief weinig leerlingen in de leeftijd van 12-14 jaar op de basisschool zitten komt doordat deze leerlingen voornamelijk in de eerste helft van het schooljaar zijn getest terwijl de teldatum voor het CBS 1 januari is. De verdeling van de leerlingen van 12-14 jaar is lastig te beoordelen door de grote categorie "gemengd". In de steekproef zijn dit vnl. leerlingen in havo/vwo brugklassen. Bij de leeftijd van 15-16 jaar zijn er nauwelijks gemengde klassen meer en komt de steekproef goed overeen met de CBS gegevens. Het percentage leerlingen met LWOO indicatie (leerweg ondersteunend onderwijs) komt eveneens goed overeen met de landelijke cijfers.

Tabel 4

Gemiddelde ruwe somscore per land en leeftijdsgroep en gemiddelde exacte leeftijd

Leeftijdsgroep	Nederland		Duitsland		gecombineerd			
	N	gem.	N	gem.	N	gem.	sd	exacte leeftijd gem.
6 jaar	62	7.1	54	7.2	116	7.2	1.6	6.50
7 jaar	64	8.2	61	8.5	125	8.4	2.1	7.51
8 jaar	64	10.0	101	9.8	165	9.9	2.1	8.49
9 jaar	64	10.9	107	10.8	171	10.8	2.0	9.49
10 jaar	65	12.0	75	12.0	140	12.0	2.4	10.47
11 jaar	64	12.9	62	12.5	126	12.7	2.7	11.54
12 jaar	65	12.9	71	13.1	136	13.0	2.9	12.50
13 jaar	58	13.4	55	13.8	113	13.6	2.6	13.52
14 jaar	61	14.2	62	14.2	123	14.2	2.9	14.50
15 jaar	63	14.7	45	14.8	108	14.7	2.5	15.49
16-17 jaar	66	15.0	52	15.2	118	15.1	2.7	16.77
18-19 jaar	64	15.3	43	15.8	107	15.5	3.0	18.74
20-21 jaar	36	16.0	45	16.1	81	16.1	2.2	20.84
22-24 jaar	31	15.9	46	16.4	77	16.2	2.1	23.08
25-29 jaar	35	16.3	48	16.1	83	16.2	2.3	27.02
30-34 jaar	40	16.3	29	16.1	69	16.2	2.4	32.37
35-40 jaar	36	15.9	39	16.1	75	16.0	2.6	38.00
Totaal	938	13.0	995	12.8	1933	12.9	3.7	15.05

Ofschoon weigering op het niveau van gemeenten, scholen en deelnemers de uitvoering van het onderzoek heeft bemoeilijkt, geeft de controle en zo nodig bijstelling van de normgroep die is uitgevoerd met betrekking tot achtergrondgegevens, voldoende vertrouwen in de representativiteit van de steekproef.

In Duitsland is gelijktijdig met het Nederlandse onderzoek een normeringsonderzoek gedaan met een vergelijkbare opzet. In Tabel 4 wordt een overzicht gegeven van de aantallen per land en leeftijdsgroep, van de gemiddelde leeftijd en van de ruwe somscore (de subtestscores zijn eerst gedeeld door de standaarddeviatie). Met een regressieanalyse met leeftijd, kwadraat van de leeftijd, land en interactie tussen land en leeftijdstermen als onafhankelijke variabelen, is onderzocht of land en de interactie van land met leeftijd significant ($p < .01$) samenhangt met de subtestscores en met de ruwe totaalscore. Bij geen van de analyses bleek het hoofdeffect voor land of het interactie-effect tussen land en leeftijd significant. Op grond hiervan zijn de normen berekend over de gecombineerde data van Nederland en Duitsland ($N=1933$). Uit de analyses

bleek ook dat de scores vanaf ongeveer 20 jaar vrijwel constant blijven. Daarom zijn bij de normering de leeftijdsgroepen vanaf 22 jaar samengenomen en zijn de normen voor de leeftijd van 22 tot 40 jaar gelijk.

Tabel 5
Verdelingskenmerken van de genormeerde testscores

Totale normgroep; leeftijd 6-40 jaar (N=1933)						
testscore	gem.	sd	min.	max.	scheefheid	kurtosis
Analogieën	9.99	3.00	1	19	.01	-.05
Mozaïeken	10.00	3.03	1	19	-.07	-.08
Categorieën	10.01	3.03	1	19	-.05	-.14
Patronen	9.99	3.00	1	19	-.15	-.09
SON-IQ	100.00	15.04	55	145	-.02	.01
Leeftijd 6-11 jaar (N=843)						
testscore	gem.	sd	min.	max.	scheefheid	kurtosis
Analogieën	10.05	3.01	1	19	.05	-.02
Mozaïeken	10.02	3.02	1	19	-.03	.21
Categorieën	10.02	3.01	1	19	.01	-.05
Patronen	10.07	3.09	1	19	-.01	.16
SON-IQ	100.22	15.11	55	145	.12	.26
Leeftijd 12-17 jaar (N=598)						
testscore	gem.	sd	min.	max.	scheefheid	kurtosis
Analogieën	9.96	3.03	1	19	.01	-.13
Mozaïeken	9.97	3.03	1	17	-.09	-.34
Categorieën	9.93	3.03	1	19	-.02	-.14
Patronen	9.95	3.01	1	17	-.18	-.25
SON-IQ	99.84	15.17	55	141	-.05	-.24
Leeftijd 18-40 jaar (N=492)						
testscore	gem.	sd	min.	max.	scheefheid	kurtosis
Analogieën	9.91	2.96	1	17	-.07	.02
Mozaïeken	9.99	3.07	1	16	-.12	-.25
Categorieën	10.09	3.06	1	18	-.18	-.24
Patronen	9.88	2.85	1	14	-.41	-.54
SON-IQ	99.87	14.79	55	140	-.21	-.15

De normering is uitgevoerd met het continue normeringsmodel dat in het kader van de constructie van de SON-R 5½-17 door ons is ontwikkeld. Bij dit model wordt de volledige scoreverdeling over alle leeftijden voor alle data gelijktijdig geschat als functie van (machten van de) de score, (machten van) de leeftijd en interactie tussen leeftijd en score. Aan het model kunnen restricties worden opgelegd als deze a priori vanzelfsprekend zijn. Dit model is sindsdien met verschillende variaties bij diverse tests in Nederland toegepast (bijv. GOS, KAIT, WISC-III, WNV). Bij de normering van de SON-R 2½-7 is het model nog op een wezenlijk punt aangepast. Bij het schatten van de parameters is niet langer de cumulatieve scoreverdeling van een leeftijdsgroep de afhankelijke variabele, maar de scores van iedere persoon afzonderlijk. Dit is mogelijk door de ruwe score te digitaliseren en per persoon voor elke mogelijke score een 0 of 1 te geven op grond van de behaalde score. De parameters van deze verdelingen (die samen beschouwd kunnen worden als de cumulatieve proportionele verdeling in de populatie) kunnen vervolgens met logistische regressie worden geschat als functie van score, leeftijd en interactie tussen beide. Bij de stapsgewijze uitgevoerde analyses bleken per subtest 8 tot 11 parameters voldoende om de verdelingen goed te kunnen beschrijven. Vervolgens zijn de covarianties tussen de subtests als functie van leeftijd gefit. Op grond hiervan kan ook de genormeerde totaalscore worden berekend. In Tabel 5 zijn voor drie leeftijdsgroepen enkele kenmerken van de genormeerde scores weergegeven. Het gemiddelde en spreiding komen goed met de verwachte waarden overeen. Wel beginnen de subtests, en dan met name Patronen, in de oudste leeftijdsgroep plafond-effecten te vertonen. Dit heeft ook effect op de scheefheid en kurtosis van de genormaliseerde verdelingen.

5. *Betrouwbaarheid*

De betrouwbaarheid is per subtest voor de leeftijdsgroepen tot 20 jaar en voor de gecombineerde groep vanaf 22 jaar geschat met coëfficiënt alpha. Weliswaar geeft coëfficiënt alpha een onderschatting van de betrouwbaarheid en is wat dat betreft lambda² beter, maar daar staat tegenover dat als gevolg van onderlinge afhankelijkheid van de item-scores de betrouwbaarheid juist wordt overschat. Vervolgens zijn de berekende waarden gefit als functie van leeftijd. De betrouwbaarheid van de genormeerde totaalscore, het SON-IQ, is berekend met de formule voor gestratificeerde alpha. De generaliseerbaarheid, die aangeeft hoe men op grond van deze selectie van subtests uitspraken kan doen over het domein van vergelijkbare tests, is voor de berekening niet afhankelijk van de betrouwbaarheid. De generaliseerbaarheid wordt namelijk gebaseerd op de correlaties tussen de subtests.

De waarden van de betrouwbaarheid en generaliseerbaarheid staan in Tabel 6. De betrouwbaarheid van de subtests is hoog, gemiddeld .87. De betrouwbaarheid is bij de oudere kinderen hoger dan bij de jongste leeftijdsgroep, bij de volwassen neemt de betrouwbaarheid van de subtests iets af. De verschillen tussen subtests en leeftijdsgroepen zijn echter gering. De betrouwbaarheid van het SON-IQ is .94 bij de jongste leeftijdsgroep en .96 in de oudere groep en bij de volwassenen. De generaliseerbaarheid neemt toe van .82 naar .86, en is .85 in de oudste groep. Gemiddeld is deze .84. Ter vergelijking: de generaliseerbaarheid van de SON-R 5½-17 is .85 en van de verkorte versie .77. De generaliseerbaarheid van de SON-R 2½-7 is gemiddeld .78.

In Nederland en Duitsland is in combinatie met het normeringsonderzoek bij 116 personen van verschillende leeftijden met een interval van gemiddeld vier maanden de SON-R 6-40 voor een tweede keer afgenomen. De stabiliteit van de subtests is gemiddeld .79 en relatief laag voor Analogieën en hoog voor Patronen (zie Tabel 6). Voor het SON-IQ is de correlatie .92 (de correlaties zijn gecorrigeerd voor restriction of range). Het gemiddelde bij de eerste afname was voor deze groep personen 104.4 (sd=15.7). Bij de tweede afname was het gemiddelde 109.6 (sd=17.3). Er is dus sprake van een duidelijk leereffect.

6. *Begripsvaliditeit*

Onder begripsvaliditeit wordt verstaan onderzoek naar de betekenis van de testuitkomst en de beoordeling of de test de eigenschap meet die wordt verondersteld. Met de analyse van de validiteitsgegevens van de SON-R 6-40 is alleen een begin gemaakt. We beperken ons hier tot de

vraag naar de interne structuur van de test en gegevens over de samenhang van het SON-IQ van de SON-R 6-40 met andere intelligentietests.

Tabel 6

Betrouwbaarheid, generaliseerbaarheid en stabiliteit van de genormeerde testscores

	betrouwbaarheid					generaliseerbaarheid
	Ana	Moz	Cat	Pat	SON-IQ	SON-IQ
Leeftijd 6-11 jaar	.85	.84	.86	.87	.94	.82
Leeftijd 12-17 jaar	.89	.87	.90	.90	.96	.86
Leeftijd 18-40 jaar	.89	.85	.89	.91	.96	.85
Totaal (6-40 jaar)	.87	.85	.88	.89	.95	.84
Hertest (N=116)	.72	.82	.77	.84	.92	

Gemiddeld zijn de correlaties tussen de subtests voor de totale normgroep .56 (tabel 7). De correlaties nemen eerst met de leeftijd toe en nemen vervolgens enigszins af bij de volwassenen. De correlaties zijn hoog en geven aan dat de onderdelen van de test in hoge mate een gemeenschappelijk concept meten. Aanvankelijk was de opzet om voor de performale onderdelen (Mozaïeken en Patronen) en de redeneertests (Categorieën en Analogieën) aparte schaalscores te berekenen, evenals dat bij de SON-R 2½-7 is gedaan. Hoewel de correlatie tussen Mozaïeken en Patronen met .67 de hoogste is en Categorieën de hoogste correlatie heeft met Analogieën, is de correlatie van Analogieën met de perceptuele tests zo hoog dat een tweedeling minder voor de hand ligt. Ook bij de SON-R 5½-17 was al gebleken dat Analogieën zowel een sterke ruimtelijke als een redeneertest is, hetgeen bijdraagt aan de hoge validiteit van deze subtest.

Een principale componenten analyse met twee componenten toont dat de vier subtests van de SON-R 6-40 alle hoog laden op de eerste ongeroteerde algemene component (gemiddeld is de lading .92) en dat de tweede component weinig toevoegt (zie Tabel 8). De eigenwaarde van de tweede component is met .28 al bijzonder laag.

Tabel 7

Correlaties tussen de genormeerde subtestscores

	Cat-Ana	Cat-Moz	Cat-Pat	Ana-Moz	Ana-Pat	Moz-Pat
Leeftijd 6-11 jaar	.45	.40	.45	.56	.53	.66
Leeftijd 12-17 jaar	.61	.56	.59	.64	.65	.70
Leeftijd 18-40 jaar	.58	.53	.52	.61	.63	.66
Totaal (6-40 jaar)	.53	.48	.51	.59	.59	.67

Tabel 8*Principale componenten-ansalyse*

Eigenwaarde	ladingen – niet geroteerd				ladingen – varimax rotatie			
		Comp. 1	Comp. 2		Comp. 1	Comp. 2		Comp. 2
Comp. 1	3.38	Ana	.92	-.04	Ana	.76		.53
Comp. 2	.28	Moz	.93	-.23	Moz	.88		.38
Comp. 3	.20	Cat	.89	.45	Cat	.44		.89
Comp. 4	.13	Pat	.94	-.16	Pat	.84		.45

Voor een intelligentietest is de mate van samenhang met andere intelligentietests een belangrijk aspect van de begripsvaliditeit. In Tabel 9 wordt een overzicht gegeven van het SON-IQ met een aantal totaalscores van andere intelligentietests. De WISC, de WAIS en de WNV zijn afgenomen in het kader van het normeringsonderzoek. Met een tussenperiode van enkele maanden is zowel in Nederland als Duitsland een andere test afgenomen. Voor de WAIS en de WNV worden de gecombineerde uitkomsten gepresenteerd. Voor de WISC is dit lastig omdat in Nederland de WISC-III wordt gebruikt en in Duitsland de WISC-IV. Daarom worden eerst de Nederlandse uitkomsten gegeven en vervolgens de gecombineerde uitkomsten voor die scores die vergelijkbaar zijn. De NIO-gegevens zijn afkomstig van scholen waar deelnemers aan het normeringsonderzoek zijn getest. De NIO is eerder afgenomen dan de SON-R 6-40, soms één tot twee jaar. Alle correlaties zijn gecorrigeerd voor 'restriction of range'.

Tabel 9*Correlaties van het SON-IQ met criterium tests (correlaties gecorrigeerd voor restriction of range)*

Test	Onderdeel	Land	N	r
WISC-III	totaal IQ	NL	36	.74
	verbaal IQ	NL	36	.55
	performaal IQ	NL	36	.81
WISC-III/IV	totaal IQ	NL+DE	71	.75
	vb-factor	NL+DE	71	.62
	vs-factor	NL+DE	71	.48
WAIS-III	totaal IQ	NL+DE	50	.83
	verbaal IQ	NL+DE	50	.74
	performaal IQ	NL+DE	50	.81
WNV	totaal score	NL+DE	123	.78
NIO	totaal	NL	74	.77
	verbaal	NL	74	.59
	symbolisch	NL	74	.76

NL=Nederland; DE=Duitsland

De gemiddelde correlatie van het SON-IQ met het totaal-IQ op de WISC, WAIS, WNV en de NIO is .78. De hoogste correlatie wordt gevonden met de WAIS (.83) en de laagste met de WISC (.75). Met het performale/symbolische deel van WISC, WAIS en NIO is de correlatie gemiddeld .79 en met het verbale deel .63. Dit betekent dat ondanks het niet-verbale karakter van de SON-R 6-40 het SON-IQ toch een zeer duidelijke samenhang vertoont met verbale intelligentie.

Tabel 10

Samenhang van SON-IQ met onderwijsniveau voor leeftijd 12-17 jaar (N=313)

Niveau	N	%	SON-IQ	
			gem.	sd
Basisonderwijs	19	6.1%	96.0	11.2
S(B)O en PRO	18	5.8%	77.8	15.2
Voortgezet onderwijs				
Vmbo bbl/kbl	76	24.3%	89.9	10.0
Vmbo (g)tl	56	17.9%	98.1	11.4
Havo	43	13.7%	106.0	10.9
Vwo	51	16.3%	114.6	12.1
gemengd – vmbo	7	2.2%	82.6	3.6
gemengd - vmbo/havo	7	2.2%	100.0	12.5
gemengd – havo/vwo	21	6.7%	110.4	10.5
MBO	15	4.8%	101.4	8.3

7. Criteriumvaliditeit

Aangezien verondersteld kan worden dat intelligentie een factor is die sterk samenhangt met leervermogen en schoolsucces, is de samenhang tussen de IQ-score en schoolprestaties en onderwijsniveau een belangrijke indicator van de criteriumvaliditeit. Op dit gebied zijn bij het normeringsonderzoek veel gegevens verzameld die nog nader moeten worden geanalyseerd. De samenhang van het SON-IQ met het onderwijsniveau bij de leerlingen in de leeftijd van 12-17 jaar geeft echter een goede indicatie van de criteriumvaliditeit (Tabel 10). De gemiddelde scores nemen volgens verwachting toe, en eta, de maat voor niet-lineaire samenhang, is .69.

Flynn-effect

Op dit moment loopt nog een onderzoek bij 100 leerlingen in groep 5, 6 en 7 van een basisschool, waarbij of de SON-R 6-40, of de corresponderende subtests van de SON-R 5½-17 worden afgenomen. Na enkele maanden maakt iedere leerling de alternatieve test. Dit zal het mogelijk maken in te schatten in hoeverre de normen van de SON-R 5½-17 zijn verouderd. Op dit moment is de eerste fase afgerond. Het gemiddelde IQ op de SON-R 5½-17 is 107.8 (N=50); het gemiddelde IQ op de SON-R 6-40 is 102.9 (N=50). Dit verschil van 5 IQ-punten is minder groot dan het verschil van 8 à 9 punten dat op grond van het Flynn-effect door ons werd verwacht.

Een andere indicatie dat de normen niet zo sterk zijn verouderd, blijkt uit de vergelijking van de IQ-scores van de SON-R 5½-17 met de WNV zoals die in de technische handleiding van de WNV worden vermeld. Bij 167 personen waarbij zowel de WNV als (in het overgrote deel van de gevallen) de SON-R 5½-17 is afgenomen, is het gemiddelde IQ bij de WNV 86.1 en bij de andere

test is het gemiddelde 85.6. Dat er geen significant verschil tussen de gemiddelden kan worden vastgesteld, is opmerkelijk. Het betekent dat de normen van de SON-R 5½-17 niet zijn verouderd, ofwel dat de normen van de WNV te hoog uitvallen. Het laatste is meer aannemelijk en zou betekenen dat de normen van de WNV niet juist zijn, evenals dit eerder bij de WAIS-III en de WISC-III het geval is geweest.

Conclusie

De SON-R 6-40 is een goede indicator voor algemene intelligentie, die op niet-verbale wijze kan worden afgenomen en die voor een breed leeftijdsbereik kan worden ingezet. De opzet en uitvoering van het normeringsonderzoek geeft vertrouwen in de normen en het gebruik van het computerprogramma maakt exacte berekening van verschillende typen scores mogelijk. De beperking tot vier subtests heeft door de uitbreiding van het aantal items in vergelijking tot de SON-R 5½-17 niet geleid tot een wezenlijke verlaging van betrouwbaarheid en generaliseerbaarheid. De subtests differentiëren nu beter, vooral voor de lagere niveaus. De IQ-score is stabiel en de indicaties voor de validiteit van de test zijn zeer positief. De beperking van de afnameduur tot één uur maakt het ook gemakkelijker om de test als aanvulling of als 'second opinion' te gebruiken in situaties waarin een verbale test is afgenomen.