



# RAKIT-2

PSYCHOMETRISCHE EIGENSCHAPPEN  
DEEL 2

White paper

# RAKIT-2

Psychometrische eigenschappen  
Deel 2

S. Pieters MSc

Testadvies, Pearson Assessment and Information B.V.

J.E. Dek MSc

Product Developer Educational Psychology, Pearson Assessment and Information B.V.

Drs. A.P. Kooij

Uitgever, Pearson Assessment and Information B.V.

PEARSON

# Inhoud

	<b>Samenvatting</b>	4
1	<b>Inleiding</b>	5
2	<b>RAKIT-2: Normeringssteekproef</b>	6
3	<b>Betrouwbaarheid: Interne consistentie</b>	7
4	<b>Betrouwbaarheid: Test-Herteststabiliteit</b>	8
5	<b>Factoranalyse</b>	9
6	<b>Neventests</b>	10
7	<b>Criteriumvaliditeit</b>	12
8	<b>Referenties</b>	14

# Samenvatting

In 2012 is de RAKIT-2 op de markt gekomen, de herziene editie van de RAKIT. Naar deze herziene versie is een uitgebreid onderzoek gedaan: er is data verzameld om nieuwe normen te genereren en er is betrouwbaarheids- en validiteitsonderzoek gedaan. In deze white paper worden de resultaten van deze onderzoeken beschreven.

Het normeringsonderzoek voor de RAKIT-2 heeft plaatsgevonden onder 1272 kinderen uit Nederland en Vlaanderen. Op basis van deze gestratificeerde steekproef zijn dagnormen gecreëerd op basis waarvan de score van elk kind omgezet kan worden naar een standaardscore.

Bij het onderzoek naar de betrouwbaarheid van de RAKIT-2 werd zowel naar cronbachs  $\alpha$  als  $\lambda^2$  gekeken. De betrouwbaarheden van de subtests waren over het algemeen .80 of hoger. De betrouwbaarheden van het Totaal IQ en het Verkort IQ liggen allen boven de .90. Ook is er onderzoek gedaan naar de test-herteststabiliteit. Voor de subtests worden gecorrigeerde correlaties gevonden van .71 - .85. Voor de factoren en IQ-scores worden waarden gevonden van .68 - .89.

Om de structuur van de subtests van de RAKIT-2 te onderzoeken werd een confirmatieve factoranalyse gedaan. Uit deze factoranalyse blijkt dat een model met vier eerste orde factoren en één tweede orde factor het best ondersteund te worden. Dit betekent dat er sprake is van vier factoren; een perceptuele redeneerfactor, verbale leerfactor, ruimtelijke oriëntatie en tempo-factor, verbale vlotheidsfactor maar ook van een sterke algemene g-factor; het totaal IQ. Deze factorindeling is voor jongere en oudere kinderen gelijk, wat een verbetering is ten opzichte van de RAKIT.

Om de congruente validiteit te bekijken is de RAKIT-2 met zowel de WPPSI-III-NL als de WISC-III<sup>NL</sup> vergeleken. Voor kinderen van 6 jaar en ouder, waarbij naast de RAKIT-2 de WISC-III<sup>NL</sup> is afgenomen, is de congruente validiteit goed te noemen. Voor jongere kinderen lijken de RAKIT-2 en de WPPSI-III-NL meer individuele variantie te hebben. Maar toch zijn de gemiddelde IQ- en verkort IQ-scores voor beide tests vrijwel gelijk.

Voor de criteriumvaliditeit is er gekeken naar kinderen met eerder vastgestelde lage IQ-scores. Hieruit bleek dat de RAKIT-2 goed differentieert, ook bij kinderen met lagere IQ-scores. Ook werden de schoolprestaties van een groep kinderen gecorreleerd met de RAKIT-2 scores. De RAKIT-2 laat een goede samenhang zien met de schoolprestaties van deze kinderen. Al dit onderzoek vormt gedegen ondersteuning voor de betrouwbaarheid en validiteit van de RAKIT-2.

# 1 Inleiding

In dit whitepaper wordt een overzicht gegeven van het normeringsonderzoek dat gedaan is naar de RAKIT-2 en op basis waarvan de normen berekend zijn. Ook worden het onderzoek en de resultaten van het validiteits- en betrouwbaarheidsonderzoek besproken. In de tekst worden ook enkele aanbevelingen gegeven hoe om te gaan met de gevonden resultaten. Voor meer informatie verwijzen wij u graag naar de handleiding van de RAKIT-2.

# 2 RAKIT-2 Normeringssteekproef

## *Aparte normen Nederland en Vlaanderen*

Voor de normering van de RAKIT-2 zijn tussen september 2009 en november 2011 in totaal 1272 afnames gedaan bij 850 kinderen in Nederland en 422 in Vlaanderen in de leeftijd van 4.0 t/m 12.5 jaar. In de vorige RAKIT waren geen Vlaamse kinderen meegenomen.

Tijdens het onderzoek werden bij een aantal subtests verschillen gevonden in de gemiddelde ruwe scores tussen Nederlandse en Vlaamse kinderen. De verschillen tussen de gemiddelde subtestscores van Nederlandse en Vlaamse kinderen zijn voor de meeste subtests klein. Voor de subtests Analogieën, Kwantiteit en Schijven zijn wel significante verschillen gevonden, hetgeen tot de beslissing heeft geleid aparte normen voor Nederland en Vlaanderen te construeren.

## *Scoring en rapportage*

Bij het verzamelen van de normen van de RAKIT-2 is gezorgd dat per leeftijd, sekse, regio, urbanisatie, etniciteit, onderwijsvorm en opleidingsniveau moeder een zo goed mogelijke afspiegeling van de doelpopulatie getrokken is. Op basis van deze data zijn er dagnormen gecreëerd die niet in de handleiding zijn opgenomen, omdat deze uit teveel tabellen zouden bestaan. Deze normen zijn daarom opgenomen in ons digitale scoringsprogramma p2O. In dit scoringsprogramma hoeven alleen de ruwe scores van een afname ingevoerd te worden. Alle berekeningen worden vervolgens door het scoringsprogramma gedaan en verwerkt in een rapportage. In de rapportage worden de ruwe subtestscores met bijbehorende standaardscores gegeven.

## *Geschiktheid voor benedengemiddeld IQ-niveau*

De range van het IQ loopt in de RAKIT-2 van 40 t/m 145. De ondergrens van 40 die bij de RAKIT-2 gehanteerd wordt, ligt lager dan bij de meeste IQ-instrumenten. Dit komt omdat bij de RAKIT-2 extra makkelijke items zijn opgenomen en er bij de normering ook extra aandacht is besteed bij kinderen met lagere capaciteiten. Het instrument is dus goed te gebruiken bij deze doelgroep. Er is in de RAKIT-2 ook een mogelijkheid tot het berekenen van een Verkort IQ. Dit Verkort IQ is gebaseerd op zes subtests, waardoor bij kinderen met problemen zoals motorische problemen of een korte concentratiespanne toch een IQ-score verkregen kan worden.

## *Dagnormen*

Er is gekozen voor dagnormen, omdat deze vorm van normen gecreëerd kan worden op basis van continue normering die bij deze test is toegepast en dit de meest precieze normen zijn. Deze normen vergelijken elk kind zo eerlijk mogelijk met zijn leeftijdsgenoten.

# 3 Betrouwbaarheid: Interne consistentie

Over alle leeftijden heen liggen de inter-itemrelaties (Chronbach's alpha) van de verschillende factorscores voor de Nederlandse steekproef tussen .83 en .95 (zie tabel 1). Voor de Vlaamse steekproef liggen deze inter-itemrelaties tussen .79 en .95. De inter-itemrelatie-indices van het Verkort en Totaal IQ liggen tussen .89 en .97 voor Nederland, en tussen .87 en .97 voor Vlaanderen. Deze waarden zijn zonder meer goed te noemen.

Tabel 1

**Gemiddelde Coëfficiënten  $\alpha$  (inter-itemrelaties) per subtest, factor, Totaal IQ en Verkort IQ, voor Nederland en Vlaanderen**

Subtest/factor	Nederland	Vlaanderen
Figuur Herkennen	0.87	0.86
Exclusie	0.88	0.88
Geheugenspan	0.84	0.85
Woordbetekenis	0.85	0.86
Doolhoven	0.86	0.84
Analogieën	0.89	0.88
Kwantiteit	0.91	0.91
Schijven	0.70	0.69
Namen Leren	0.84	0.77
Verborgene Figuren	0.90	0.91
Ideeënproductie	0.82	0.75
Vertelplaat	0.72	0.67
<b>Totaal IQ</b>	<b>0.96</b>	<b>0.95</b>
<b>Verkort IQ</b>	<b>0.93</b>	<b>0.91</b>
Perceptuele redeneerfactor	0.95	0.95
Verbale leerfactor	0.89	0.86
Ruimtelijke oriëntatie en tempofactor	0.84	0.82
Verbale vlotheidsfactor	0.83	0.79

Naast de coëfficiënt  $\alpha$  (inter-itemrelaties) zijn de lambda2-indices (Guttman, 1945) per subtest, factor en Totaal en Verkort IQ, per leeftijdsband en over de leeftijd heen berekend. De lambda2-waarden komen alle hoger uit dan de eerder genoemde inter-itemrelaties, al is het maximale verschil met de coëfficiënten  $\alpha$  0.03. Op grond van zowel de  $\alpha$ - als de lambda2-coëfficiënten kunnen we zeggen dat de RAKIT-2 een betrouwbare test is met betrouwbare onderdelen, zeker ook als het gaat om de factor- en totaalscores. Dit betekent dat met de RAKIT-2 de gemeten waarde dicht in de buurt komt van de daadwerkelijke waarde en dat daar niet teveel fluctuatie in zit. Ook geeft dit aan dat de scores die met de RAKIT-2 gemeten worden grotendeels vrij zijn van toevallige fouten.

# 4 Betrouwbaarheid: Test-herteststabiliteit

Een ander aspect van de betrouwbaarheid van een test is de mate waarin een test bij herhaalde afname hetzelfde meet. Om dit te onderzoeken is de RAKIT-2 tweemaal afgenomen bij een groep van 118 kinderen. Deze steekproef is aselekt getrokken uit twee leeftijdsgroepen uit de beide normeringssteekproeven (Nederland en Vlaanderen). Er is gekeken naar een groep van 6- en 7-jarige kinderen en een groep van 9- en 10-jarige kinderen.

Over beide leeftijdsgroepen heen zijn de test-hertestcoëfficiënten alle boven .80. Ook op subtestniveau is de test-hertestbetrouwbaarheid van de RAKIT-2 goed te noemen, op twee indices na liggen alle waarden (ruim) boven .70. De Verbale vlotheidsfactor kent relatief de laagste toenames, terwijl de Verbale leerfactor (de naam zegt het in feite al) de grootste gemiddelde vooruitgang laat zien van bijna één standaarddeviatie, vooral als gevolg van het herhaald aanbieden van de subtest Namen Leren. Dit is niet verwonderlijk; immers het doel van deze subtest is dat kinderen zich de namen van de poezen en vlinders bij de plaatjes tijdens de afname blijven herinneren. Blijkbaar blijft deze herinnering langer hangen dan de duur van de afname, ondanks de tussenperiode. Voor de jongste kinderen laat ook de subtest Doolhoven een aanzienlijke vooruitgang zien na een tweede testafname. Er dient dan ook in de praktijk een aanzienlijke tijdsperiode tussen 2 afnames gehanteerd te worden, gezien het verschil in Totaal IQ van 10 punten tussen de 2 afnames.



# 5 Factoranalyse

Om te toetsen of de RAKIT-2-data in voldoende mate overeenkomen (*fit*) met het factormodel van de oude RAKIT werden er confirmatieve analyses uitgevoerd.

De confirmatieve analyses ondersteunen de indeling van de RAKIT-2 subtests in vier duidelijk te onderscheiden cognitieve domeinen en maken duidelijk dat de structuur van de RAKIT-2 over de tijd geen wezenlijke veranderingen heeft ondergaan. Wel blijkt dat voor de jongste groep de vroegere indeling een gelijke fit oplevert als de indeling van de oudere leeftijdsgroep. Daarom is er besloten tot een factorstructuur met vier factoren die hetzelfde is voor alle leeftijdsbanden in Nederland en Vlaanderen. Dit model ondersteunt zowel de data als de theorie het best. Dat deze indeling hetzelfde is voor alle leeftijdsbanden is een voordeel ten opzichte van de oude RAKIT, waarbij voor de jonge en oudere kinderen een verschillende factorindeling gehanteerd werd. Dit leidde in het verleden mogelijk tot onduidelijkheid.

# 6 Neventests

De congruente of soortgenootvaliditeit van een test kan worden bepaald door de samenhang van de betreffende test met een andere, gelijksoortige test vast te stellen. Ter bepaling van de soortgenootvaliditeit van de RAKIT-2 is bij 133 kinderen naast de RAKIT-2 ofwel een WPPSI-III-NL ofwel een WISC-III<sup>NL</sup> afgenomen.

## *Vergelijking tussen de RAKIT-2 en de WPPSI-III-NL*

Van de WPPSI-III-NL werden de kernbatterij en de subtests Onvolledige Tekeningen en Receptieve Woordenschat afgenomen. Het gemiddeld Totaal IQ in deze groep van de RAKIT-2 bedroeg 104 en het Verkort IQ 102. Het gemiddeld WPPSI-III-NL IQ bedroeg 104, het Verbaal IQ was 105 en het Performaal IQ 105. De correlatie tussen de beide Totaal IQ's is .76. De samenhang tussen de factor Perceptueel redeneren (RAKIT-2) en het TIQ van de WPPSI-III-NL is .73, tussen Verbaal leren (RAKIT-2) en het TIQ van de WPPSI-III-NL is .59. De factoren Ruimtelijke oriëntatie en tempo en Verbale vlotheid meten vaardigheden die veel minder voorkomen in de WPPSI-III-NL waardoor de correlaties lager liggen. De intercorrelaties tussen de subtests laten ook een wisselend patroon zien en variëren van .76 (Analogieën en Woordenschat) en .66 (Namen Leren en Informatie, Exclusie en Matrix Redeneren) tot over de hele linie lage correlaties tussen Schijven, Ideeënproductie en Vertelplaat enerzijds en WPPSI-III-NL subtests anderzijds.

## *Vergelijking tussen de RAKIT-2 en de WISC-III-NL*

Bij 67 kinderen werd naast de RAKIT-2 ook de WISC-III<sup>NL</sup> afgenomen. Van de WISC-III<sup>NL</sup> zijn alle kernsubtests en de subtest Doolhoven afgenomen.

Het gemiddeld Totaal IQ bedroeg in deze studie 106, Verkort IQ ook 106. Het gemiddeld WISC-III<sup>NL</sup> IQ bedroeg 106, het Verbaal IQ was 106 en Performaal IQ 104. De correlatie tussen de beide Totaal IQ's is .82 en tussen de verkorte versie van de RAKIT-2 en het TIQ van de WISC-III<sup>NL</sup> .80. De RAKIT-2-factor Perceptueel redeneren en het TIQ van de WISC-III<sup>NL</sup> correleren .74, tussen Verbaal Leren en het TIQ van de WISC-III<sup>NL</sup> is deze correlatie .70, tussen Ruimtelijke oriëntatie en tempo en het TIQ van de WISC-III<sup>NL</sup> .57 en tussen Verbale vlotheid en het TIQ van de WISC-III<sup>NL</sup> .48. De intercorrelaties tussen de subtests zijn sterker dan bij de WPPSI-III-NL en variëren van .80 tussen Schijven en Figuur Leggen, .79 tussen Exclusie en Doolhoven tot lage correlaties tussen onder andere Doolhoven en Woordenschat en Figuur Herkennen en Begrijpen.

## *Conclusies over neventests*

De congruente validiteit van de RAKIT-2 is goed te noemen, zeker voor de leeftijdsgroepen van 6 jaar en ouder. Voor de jongere kinderen lijken de RAKIT-2 en WPPSI-III-NL iets meer individuele variantie te hebben, maar de gemiddelde IQ- en verkort IQ-scores zijn voor beide tests vrijwel gelijk.

Hieruit blijkt dat de RAKIT-2 een groot deel overlap heeft met de beide Wechslers, maar ook een deel unieke vaardigheden meet. Hierdoor kan de RAKIT-2 goed ingezet worden als hertest; om korte tijd na een afname van een Wechsler een nieuwe IQ-test te doen voor herevaluatie doeleinden. Ook kunnen losse subtests van de RAKIT-2 gebruikt worden om eerder verkregen scores te evalueren.

Bij het kiezen van een intelligentietest is, verdient het in een aantal gevallen de voorkeur om voor de RAKIT-2 te kiezen:

- als een lagere begaafdheid vermoed wordt; omdat de RAKIT-2 een IQ-range heeft van 40-145 is het mogelijk ook bij kinderen met een lagere begaafdheid op valide wijze het IQ in kaart te brengen.
- Bij leerlingen met fysieke of psychische handicaps. De RAKIT-2 biedt een verkorte vorm die specifiek geschikt is om in korte tijd (45-75 minuten) het totaal IQ in kaart te brengen. Deze verkorte versie doet slechts een gering beroep op de motoriek en de subtests van deze verkorte versie zijn door experts aangeduid als bruikbaar bij kinderen met stoornissen.
- als er korte tijd geleden een WPPSI-III-NL, WISC-III<sup>NL</sup> of SON-R is afgenomen, vormt de RAKIT-2 een goed instrument om een hertest- of evaluatiemeting mee uit te voeren.
- als er een vergelijking gewenst is tussen prestaties doorlopend op de basisschool; omdat de RAKIT-2 specifiek geschikt is voor de hele basisschoolleeftijd, is het mogelijk om bijvoorbeeld een vergelijking te maken van de prestaties in de kleuterklas en in groep 8.

# 7 Criteriumvaliditeit

## *Speciale groepen lager begaafde kinderen*

Ten tijde van het normeringsonderzoek van de RAKIT-2 werd de vernieuwde test ook afgenomen bij drie speciale groepen kinderen. Omdat een van de doelen van de herziening van de RAKIT was het kunnen testen van kinderen met een relatief laag intelligentieniveau, werd de RAKIT-2 bij een aantal kinderen uit zowel het speciaal basisonderwijs als het speciaal onderwijs de RAKIT-2 afgenomen. Toetsing liet, zoals verwacht op basis van de eerder gemeten IQ's, over de hele linie significante verschillen in gemiddelden zien, tussen enerzijds de drie (lage) IQ-groepen en anderzijds de normgroep.

Het bleek mogelijk te zijn om kinderen met deze lage intelligentiescores te testen met de RAKIT-2, al wordt wel aanbevolen bij hen de verkorte versie van de test af te nemen, gezien de korte aandachtsspanne van de kinderen. Indien dit toch niet gewenst is, bijvoorbeeld omdat men een gedifferentieerd beeld van de cognitieve mogelijkheden van het kind wil krijgen, verdient testafname over meer meetmomenten de voorkeur. De verkorte vorm van de RAKIT-2 is overigens dusdanig samengesteld dat de meest toegankelijke subtests voor kinderen met lage IQ's zijn opgenomen. Deze keuze is gebaseerd op advies van experts die werken met kinderen in het speciaal onderwijs en instellingen voor kinderen met een verstandelijke beperking. Daarnaast loopt het IQ-bereik van de RAKIT-2 tot 40, wat dus zelfs in de laagste IQ-gebieden nog een gedifferentieerde score biedt.

## *Intelligentie en schoolprestatie*

Voor onderzoek naar de criteriumvaliditeit zijn per onderzochte leeftijdsband Cito data verzameld bij een groep kinderen uit de Nederlandse steekproef. Voor 202 kinderen uit de Nederlandse steekproef waren scores op diverse Cito-toetsen bekend. De correlaties tussen de Cito-toetsen en de RAKIT-2 factor- en IQ-scores is te zien in tabel 2.

Tabel 2

**Correlatie van Totaal IQ, Verkort IQ en factorscores van de RAKIT-2 met scores op Cito-toetsen voor kinderen uit leeftijdsbanden II en III**

	Technisch lezen (N= 128)	Taal (N= 91)	Spelling (N= 184)	Rekenen (N= 197)	Begrijpend luisteren (N= 45)	Begrijpend lezen (N= 167)	Woordenschat (N= 123)	Gem.	SD
Totaal IQ	0.54**	0.47**	0.45**	0.53**	0.50**	0.59**	0.46**	103.5	15.1
Verkort IQ	0.52**	0.47**	0.42**	0.51**	0.50**	0.57**	0.47**	103.6	15.1
PCR	0.50**	0.38**	0.44**	0.52**	0.45**	0.56**	0.35**	15.6	3.6
VBL	0.44**	0.54**	0.40**	0.44**	0.45**	0.61**	0.46**	15.7	4.1
ROT	0.39**	0.18	0.27**	0.28**	0.18	0.21**	0.30**	15.6	3.4
VVL	0.21*	0.28**	0.11	0.23**	0.31**	0.24**	0.21*	15.9	3.6
Gem.	4.2	4.2	3.8	4.1	4.2	4.0	4.0		
SD	1.1	1.0	1.2	1.1	0.9	1.1	1.1		

\* significant op 0.05 niveau en \*\* significant op 0.01 niveau

We zien hoge correlaties tussen alle schooltoetsen en het Totaal IQ en Verkort IQ, variërend tussen .42 en .59. Dit betekent dat de samenhang tussen Begrijpend luisteren, Begrijpend lezen, Technisch lezen en Rekenen met het Totaal IQ het hoogst is. Veelal worden in de literatuur correlaties tussen .40 en .55 gemeld (Resing & Drenth, 2007; Drenth & Sijtsma, 2006), hetgeen dus ook voor de RAKIT-2 is geconstateerd. Hieruit blijkt dus dat de scores op de RAKIT-2 goede voorspellers kunnen zijn van schoolresultaten. Dit betekent dat de RAKIT-2 ingezet kan worden bij het geven van schooladviezen.

# 8 Referenties

- Bleichrodt, N., Resing, W.C.M., Drenth, P.J.D., & Zaal, J.N. (1987). *Intelligentiemeting bij kinderen*. Amsterdam: Pearson Assessment and Information B.V.
- Centraal Bureau voor Statistiek [CBS] -Statline (2009). Bevolkingscijfers. Retrieved april 2009 from <http://www.cbs.nl>
- Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.
- Evers, A., Lucassen, W., Meijer, R., & Sijtsma, K. (2010). *COTAN beoordelingssysteem voor de kwaliteit van tests*. Amsterdam: NIP.
- Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika*, 10, 255-282.
- NPdata (2009). Retrieved april 2009 from <http://www.npdata.be>
- Resing, W.C.M., Bleichrodt, N., Drenth, P.J.D. & Zaal, J.N. (2012). *RAKIT-2 handleiding*. Amsterdam: Pearson Assessment and Information B.V.
- Sijtsma, K. (2009). Over misverstanden rond Cronbachs alfa en de wenselijkheid van alternatieven. *De Psycholoog*, 44, 561-567.
- Statistics Belgium [statbel] (2009). Retrieved April 2009 from <http://statbel.fgov.be/nl/statistieken/cijfers>
- Studiedienst van de Vlaamse regering (2009). Retrieved April 2009 from <http://www4.vlaanderen.be/dar/svr/cijfers/exceltabellen/onderwijs/kwaliteitszorg/ONDEKWAL011.xls>





Pearson Assessment and Information BV

Radarweg 60-A1, 1043 NT Amsterdam

Postbus 78, 1000 AB Amsterdam

T: +31 (0)20 581 5500

E: [info-nl@pearson.com](mailto:info-nl@pearson.com)

[www.pearsonclinical.nl](http://www.pearsonclinical.nl)

[www.pearsonclinical.be](http://www.pearsonclinical.be)

Twitter: [@PearsonNL](https://twitter.com/PearsonNL)