



Handleiding cognitieve capaciteitentest

Cijferreeksen MBO

Testconstructie

Voor deze test werd een itemset gegenereerd met vierkeuze-items volgens een vast format: een reeks getallen wordt gepresenteerd waarin een logische volgorde te ontdekken valt. De onderzochte dient vervolgens aan te geven welk van vier mogelijke getallen logischerwijs op de gepresenteerde reeks zou moeten volgen. Alle items werden vooraf door twee psychologen gescreend en in geval van onduidelijkheid of ambiguïteit van een item werd het betreffende item verwijderd. Uiteindelijk resulteerde dit in een set van 27 items. Deze werden online afgenomen bij een grote groep personen (N = 2752). In deze groep was een aanzienlijk percentage personen dat geen hoogst genoten opleidingsniveau had opgegeven. Deze personen werden om deze reden uit de normgroep verwijderd, evenals enkele personen jonger dan 16 jaar en enkele personen waarvan geen geslacht of leeftijd bekend was. De proefgroep werd hiermee gereduceerd tot 1758 personen (49% vrouwen en 51% mannen). Nadere informatie over deze proefgroep wordt vermeld onder Normering.

Schaalconstructie en betrouwbaarheid

Op de data van bovengenoemde 1758 personen werd vervolgens een itemanalyse gedaan. De p-waarden van de verschillende items werden onderzocht, en tevens werd op basis van item-rest correlaties onderzocht welke items minder geschikt waren om in de test opgenomen te worden. Er bleken in totaal drie items met een zeer matige item-rest correlaties ($< .10$) te zijn. Besloten werd deze drie items uit de test te verwijderen, waarna de uiteindelijke test uit 24 items bestaat. Hierna werden enkele betrouwbaarheidsindices berekend van de 24 items tellende test: alfa = .81 en Guttman's $\lambda^2 = .81$.

Normering

Leeftijdsverschillen en geslachtsverschillen

Met het oog op het onderzoeken van mogelijke leeftijdsverschillen met betrekking tot de testprestatie werd de groep opgesplitst in een aantal groepen op basis van leeftijd (16-17 jaar, 18-20 jaar, 21-30 jaar, 31-40 jaar, 41-50 jaar, 51-55 jaar en 56+). Inspectie van de frequenties van deze leeftijdscategorieën liet zien dat er vrij weinig personen ouder dan 55 hadden meegedaan aan het onderzoek (in totaal $N = 42$). Vanwege deze kleine N werd besloten deze groep personen te verwijderen. Hiermee werd de normgroep verder gereduceerd tot 1716 personen (leeftijd $M = 32.5$, $SD = 10.4$, range 16-55). Tussen de resterende leeftijdscategorieën werden slechts enkele verschillen groter dan 1 punt (ruwe scores) gevonden. Deze verschillen werden niet gemedieerd door eventuele verschillen in opleidingsniveau tussen de leeftijdscategorieën. Hierop werd besloten om ten behoeve van de normering de normgroep op basis van de gemiddelde ruwe scores op te splitsen in twee groepen: 16-40 jaar ($N = 1375$) en 41-55 jaar ($N = 341$). Als aanvulling hierop werd tevens de correlatie tussen leeftijd en testscore berekend: $r = -.17$. Mannen en vrouwen bleken gemiddeld amper verschillend te scoren: mannen $M = 14.5$, $SD = 5.5$, vrouwen $M = 14.1$, $SD = 5.5$. Hierop werd besloten de normering alleen voor genoemde twee leeftijdscategorieën en niet apart voor mannen en vrouwen uit te voeren.

Opleidingsniveau

Het opleidingsniveau werd gehercodeerd in vijf categorieën (Luteijn & Barelds, 2005):

1. basisonderwijs
2. lager beroepsonderwijs of gelijkwaardig
3. MAVO/MBO of gelijkwaardig
4. HAVO/VWO/HBO of gelijkwaardig
5. WO

Vervolgens werden de percentages van deze opleidingscategorieën berekend, apart voor mannen en vrouwen, en vergeleken met CBS-gegevens voor personen in de leeftijd van 16-55 jaar (Luteijn & Barelds, 2005). De percentages bleken voor mannen en vrouwen in de huidige normgroep vrijwel identiek. Er bleken, in vergelijking met de CBS-gegevens, vooral te veel hoog opgeleiden te zijn. Besloten werd om door middel van weging de normering te laten corresponderen met een met betrekking tot opleidingsniveau representatieve normgroep (CBS-gegevens). De wegingscoëfficiënten die daarvoor voor de vijf hiervoor genoemde opleidingsniveaus werden gebruikt waren respectievelijk 3.8, 2, 3.3, 1.0 en 0.5. Het gehercodeerde opleidingsniveau bleek $.22$ ($p < .001$) te correleren met de totaalscore op de test.

Decielen

Na weging van de normgroep met betrekking tot opleidingsniveau werden decielen berekend (groepen met elk 10% van de scores) voor personen van 16-40 jaar en personen van 41-55 jaar. Deze decielscores worden in de uitslag van de test automatisch berekend

Intelligentie

Naast de decielscores werden tevens schattingen van de IQ's gemaakt op basis van de testcores. Hiertoe werden (na weging van de opleidingsniveaus zoals hiervoor genoemd), de testcores eerst lineair getransformeerd [$Y = \arcsin(X)^{1/2}$], omdat de verdeling niet geheel normaal was (licht rechts scheve verdeling). Deze lineair getransformeerde scores werden vervolgens getransformeerd naar IQ-scores met $M = 100$ en $SD = 15$. Aangezien er slechts 24 mogelijke scores op deze test zijn, kunnen hieruit ook slechts 24 verschillende IQ-schattingen worden bepaald. In plaats van deze IQ-schattingen exact te rapporteren doet men er verstandig aan deze globaal te omschrijven (zie ook Luteijn & Barelds, 2005; Resing & Blok, 2002). De hier gebruikte indeling volgt grotendeels die van Luteijn en Barelds (2005; zie Tabel 1). In de uitslag van de test wordt deze omschrijving automatisch gegeven.

Tabel 1. Omschrijving van globale intelligentie op basis van testcores.

Omschrijvingen
Lage verstandelijke capaciteiten
Ruim benedengemiddelde verstandelijke capaciteiten
Benedengemiddelde verstandelijke capaciteiten (vgl. LBO-niveau)
Gemiddelde verstandelijke capaciteiten (vgl. MBO-niveau)
Bovengemiddelde verstandelijke capaciteiten (vgl. HBO-niveau)
Ruim bovengemiddelde verstandelijke capaciteiten (vgl. WO-niveau)
Goede verstandelijke capaciteiten (ruim WO-niveau)
Zeer goede verstandelijke capaciteiten (WO+ niveau)

Ten behoeve van een vriendelijke rapportage naar respondenten wordt er in het rapport niet gesproken van 'verstandelijke capaciteiten' maar van 'score'.

Men dient er bij zowel de rapportage van de decielscores als de globale intelligentieomschrijvingen rekening mee te houden dat de testcore ten gevolge van toevalligheden lager of juist hoger uit kan vallen. Met andere woorden: men kan pech of geluk hebben. Zo bedraagt de standaardmeetfout van de test bijvoorbeeld 2.4. Tengevolge daarvan dienen uitspraken over de testprestatie genuanceerd te worden. Deze meetfout kan tevens in een rapportage worden gerapporteerd.