

## *Persoonlijkheidstesten*

# Zin of onzin?

Auteur: **Danny Mullenders**

De een zweert erbij en gelooft dat ze mensen, teams en organisaties echt verder helpen. De ander vindt ze nutteloos of zelfs complete onzin: persoonlijkheidstesten. Ze zouden voorspellen wat voor type persoon iemand is, hoe een persoon zich in een specifieke situatie zal gedragen en of iemand geschikt is voor een bepaalde functie of context. In HR-land zijn behoorlijk wat testen in omloop. Van zeer wisselende kwaliteit.



De afgelopen twintig jaar nam het aanbod van deze persoonlijkheids- en andere psychologische testen een flinke vlucht. Vooral assessmentbureaus gebruiken deze testen, maar ook diverse trainings- en adviesbureaus zetten ze veelvuldig in als standaardonderdeel van verander- of leiderschapstrajecten. Ook zijn er organisaties die hun HR-medewerkers laten trainen in het afnemen en interpreteren van dit soort testen. Maar wat is de toegevoegde waarde en betrouwbaarheid van deze persoonlijkheidstesten? In dit artikel lees je waaraan een goede persoonlijkheidstest moet voldoen, hoe je die herkent, op welke valkuilen je hierbij beducht moet zijn, en een aantal suggesties om testen zo goed en zuiver mogelijk te laten ondersteunen bij belangrijke beslissingen over mensen.

### *Kortom, testen ondersteunen het maken van betere keuzes op basis van voorspelbaarheid.*

Om te beginnen: waarom zijn deze testen zo populair? Persoonlijkheid voorspelt voor een deel prestaties en welbevinden. Het maakt het begeleiden van individuen, teams en leidinggevend effectiever. Psychologische testen worden dan ook gebruikt om persoonlijkheidsprofielen, eigenschappen en voorkeuren van mensen in kaart te brengen. Vaak inclusief inschattingen over hoe iemand zich gedraagt onder bepaalde omstandigheden, zoals druk, in een bepaald type werk of positie, en eventueel welke risico's of ontwikkelpunten te verwachten

zijn in een bepaalde werkomgeving, functie of functieniveau.

Vanuit gebruikers gezien, zoals recruiters, een HR-afdeling, managers of raden van commissarissen, is de belangrijkste vraag bij het beoordelen van een test hoe goed de test hen helpt met voorspellen welke kandidaat het beste bij een functie past of andersom. Kortom, testen ondersteunen het maken van betere keuzes op basis van voorspelbaarheid.

#### **In vogelvlucht**

Waarschijnlijk ontwikkelden Chinezen al in 2200 voor Christus psychologische testen. In het westen ontwikkelden psychiaters in de achttiende en negentiende eeuw classificatiesystemen om mentale problemen te groeperen. Eind achttiende eeuw werden de eerste persoonlijkheidstesten afgenomen. Volgens de leer van de frenologie – later afgedaan als kwakzalverij – waren persoonlijkheidskenmerken af te leiden aan de vorm van de schedel.

In de negentiende eeuw kwam de empirische psychologie op. De nieuwe empirische methoden om ingewikkelde ideeën te vereenvoudigen en te standaardiseren, bij voorkeur in getallen, bleken ook toepasbaar bij de studie van de persoonlijkheid. Zo werd de eerste intelligentietest ontwikkeld: de Binet-Simon scale, die intelligentie mat onder andere aan de hand van beredenering, begrip en het kortetermijngeheugen.

Doordat in de Eerste Wereldoorlog grote groepen mensen geselecteerd moesten worden voor uiteenlopende functies, kwam de ontwikkeling van het testen in een

stroomversnelling. Zo ontstond erkenning dat de psychologische test een positieve bijdrage kon leveren in beslissingssituaties. Het Amerikaanse leger gebruikte ook persoonlijkheidstesten om te testen of soldaten ontvankelijk waren voor shellshock: de psychische gevolgen van een oorlog voor een soldaat, denk aan angststoornissen, PTSS en slapeloosheid. En om rekruten op leiderschapskwaliteiten te screenen.

In de jaren 40 en 50 van de vorige eeuw werden de testen ingezet om op een wetenschappelijke manier de basiskenmerken van de menselijke persoonlijkheid te ontdekken, zoals extraversie en openheid.

Vanaf de jaren 70 van de vorige eeuw nam de persoonlijkheidstest een vlucht. Vanaf toen dachten psychologen steeds vaker dat persoonlijkheid meetbaar is. Ze ontwierpen allerlei testen om bijvoorbeeld aspecten als cognitieve vaardigheden en emotionele intelligentie vast te leggen. Zulke testen werden weer door organisaties gebruikt om het succes van kandidaten proberen te voorspellen of om te kijken of iemand bij de organisatie past.

Vandaag de dag worden persoonlijkheidstesten vaak gebruikt om te testen of een kandidaat geschikt is voor een functie of binnen een team past. Er zijn zelfs bedrijven waarbij een persoonlijkheidstest het enige middel is om een voorselectie van kandidaten te maken. Ook wordt geprobeerd een beeld te vormen van de eventuele risico's die organisaties kunnen lopen met iemands persoonlijkheid in een functie en of iemand ontwikkelmogelijkheden op bepaalde vlakken heeft.

### Betrouwbaarheid en validiteit

Er bestaan gedegen meetinstrumenten om persoonlijkheidsstructuur en persoonlijkheidstrekken in kaart te brengen, waar veel onderzoek naar is gedaan. Denk aan het Big Five Model (BFM), het Five Factor Model (FFM) en HEXACO (HEX). Maar wat bepaalt nou de kwaliteit van een persoonlijkheidstest? De belangrijkste factoren hiervoor zijn validiteit en betrouwbaarheid. Validiteit is de mate waarin resultaten geldig zijn en overeenkomen met de werkelijkheid. Kortom: meet je wat je beoogt te meten? Betrouwbaarheid wil zeggen dat de resultaten hetzelfde zijn als je de test op dezelfde wijze nogmaals invult.

Verder moet er sprake zijn van een onderliggende psychologische theorie, ofwel een model waarop de test gebaseerd is. En de test moet uiteraard objectief zijn. Hiervoor bestaat een gestandaardiseerde manier van vragen stellen en scores van de antwoorden verwerken. Idealiter kan een test over de hele wereld gebruikt worden en werken de plaatselijke cultuur en gewoonten niet door in de testuitslagen.

In het beste geval is een persoonlijkheidstest dus valide en betrouwbaar. Regelmatig worden psychologische testen dan ook aanbevolen om hun uitstekende validiteit. 'Validiteit' is echter een breed begrip. In de wetenschap wordt een onderscheid gemaakt tussen verschillende soorten validiteit. De soort validiteit die hier essentieel is, is criteriumvaliditeit ofwel de voorspellende waarde.

Helaas blijkt juist deze soort validiteit afwezig in de testen die het meest gebruikt

worden. De makers van deze testen zijn meer commercieel geïntereerd. Ze winnen er niets bij om de validiteit en betrouwbaarheid van hun test te expliciteren. Vooral omdat die laag zijn.

Ook hebben veel persoonlijkheidstesten die al vele jaren gebruikt worden een gammell fundament. Ondanks dat er wel degelijk actuele kennis voorhanden is, zijn ze dikwijls gebouwd op gedateerde psychologische aannames en modellen. Jammer, want als je de verkeerde test gebruikt, kun je verkeerde beslissingen in je selectieproces nemen, waardoor je uiteindelijk niet de beste kandidaat in huis haalt. Daarom is het relevant dat je test zowel hoge betrouwbaarheid als validiteit oplevert.

### **Huidige praktijk**

In de wereld van de testen zien we vaak het horoscoopeffect, de verkeerde achtergrond bij de afnemer en de verkeerde inzet van de test waardoor de betrouwbaarheid laag is.

### **Ongeschikte achtergrond**

Naast de slechte kwaliteit van testen worden testresultaten vaak niet juist geïnterpreteerd. Ondanks dat HR-afdelingen en leidinggevendens dikwijls een aantal dagen getraind worden in methodieken, waarbij grif betaald wordt voor een vragenlijst plus rapportage, wil dat niet zeggen dat ze de test goed weten in te zetten. Dikwijls is de reden dat gebruikers geen arbeids- en organisatiepsychologische achtergrond hebben.

### **Verkeerd ingezet**

Veel organisaties gebruiken testen als selectie-instrument in een

### **Horoscoopeffect**

In 2020 analyseerden de Zweden Patrik Reman en Angelika Nordin een jaar lang zowat alles over elke persoonlijkheidstest die wordt gebruikt in recruitment. Hierbij richtten ze zich vooral op de wetenschappelijke validatie van testen. Hun conclusie? Bijna alle testen zijn eigenlijk net zo betrouwbaar als een horoscoop. Ze benaderen het niveau van een waarzegger die je toekomst met tarotkaarten voorspelt.

Hoe dit komt? Vooral door het zogenoemde 'Forereffect'. De Amerikaanse psycholoog Bertram Forer voerde in 1948 een persoonlijkheidstest uit bij zijn studenten. Vervolgens stuurde hij hen allemaal een en dezelfde algemene persoonlijkheidsomschrijving die zogenaamd op de persoonlijkheidstest gebaseerd was. Iedereen gaf terug dat deze omschrijving precies op hem van toepassing was. Precies het horoscoopeffect waarbij de ontvanger een heerlijk gevoel van herkenning krijgt; *confirmation bias*. Waarschijnlijk omdat we vanuit onze menselijke behoefte aan duidelijkheid, betekenisgeving en samenhang simpele begrijpelijke verhalen over onszelf aan elkaar willen vertellen.

Dit proces is waar de populariteit van dergelijke testen op drijft. Volgens de Zweden zou de Big Five dit effect overigens niet hebben.

recruitmentproces, die helemaal niet voor dat doel bedoeld zijn. Neem de Myers-Briggs Type Indicator (MBTI). Op basis van vier basispersoonlijkheidskenmerken worden mensen tot een van de zestien MBTI-persoonlijkheidstypes bestempeld. Dit klinkt misschien aannemelijk en grijpt zelfs terug op de ideeën van psychiater Carl Jung, maar voor werving en selectie is deze persoonlijkheidstest echt ongeschikt. Er zijn veel analyses die laten zien dat MBTI ineffectief is om het succes van mensen in een baan te voorspellen. De onderliggende theorie is niet evidencebased, en zit psychometrisch bepaald niet goed in elkaar.

### Kleuren en kwadranten

Er zijn inmiddels talloze testen die inzicht geven in de gedrags- en persoonlijkheidseigenschappen van mensen om persoonlijkheidstypen te bepalen. Helaas zijn de testen van beperkte kwaliteit veruit in de meerderheid.

Vaak gaat het dan over de zogenaamde kleurentesten. Het toepassen van dit 'kleurendenken' voor teams en organisaties ontstond in de jaren tachtig van de vorige eeuw. Psycholoog, arts en managementauteur Edward de Bono was een van de eersten die het kleurendenken toen introduceerde. De Bono is de bedenker van de 'zes denkhoeden'; een techniek waarbij deelnemers vanuit een verrassend perspectief (een kleur hoed) een vraagstuk oplossen.

Mensen ervaren werken met kleuren als aantrekkelijk. Kleurentesten brengen een persoonlijkheid namelijk terug tot een of enkele kleuren. Veel mensen herkennen zichzelf en hun collega's hier makkelijk in. En een gesprek voeren over kleuren is veel eenvoudiger dan praten over individuele eigenschappen. Maar kleurentesten zijn

simplificaties die eerder schijnhelderheid, hokjesdenken en weinig genuanceerde containertaal opleveren.

Deze testen sluiten uitstekend aan bij ons intuïtieve gevoel dat mensen in verschillende categorieën in te delen zijn. De rapporten zien er aantrekkelijk, kleurrijk en overzichtelijk uit en resultaten zijn eenvoudig en makkelijk te begrijpen, vaak tactvol beschreven en vooral positief geformuleerd. Niet in de laatste plaats is dat vaak een resultaat van de marketingcampagne van de verkoper.

*Deze testen sluiten uitstekend aan bij ons intuïtieve gevoel dat mensen in verschillende categorieën in te delen zijn.*

### Pseudowetenschappelijke testen

Onze behoefte aan praktische kleurinstrumenten om mensen in te kunnen delen en onszelf te kunnen begrijpen heeft geleid tot de ontwikkeling van pseudowetenschappelijke testen die mensen in kleuren kadert of in kwadranten indeelt. Dergelijke vaak korte testen, die dikwijls uit tien tot twintig vragen bestaan en pakweg vijf minuten kosten, zijn misschien leuk om even in te vullen, maar de wetenschappelijke basis van de meeste kleurentesten is uitermate dun. Meestal werken ze niet omdat de onderliggende theorie niet klopt en het onduidelijk is wat er gemeten wordt. Mensen worden ingedeeld in te simplificerende categorieën.

Waarom deze manier van kijken naar mensen in hapklare brokken in eerste instantie aantrekkelijk en dus verleidelijk

is? Wetenschappelijke modellen zijn genuanceerd en complex en kleurentesten bieden eenvoud, zijn makkelijk in gebruik en zijn positief. Bovendien worden de uitkomsten meestal als 'waar' ervaren. Ze zijn immers een bevestiging van wat de invuller of zelfs de interpreteerder vooraf zelf al enigszins verwachten. Samen biedt dit een illusie van grip en controle.

### **Risico**

Aan de simplificatie van dit soort kleurentesten zitten risico's. In eerste instantie omdat het echte duidelijkheid juist voorkomt. Zo kan een goed gesprek achterwege blijven, omdat een kleur minder zegt dan een specifiekere beschrijving van gedrag. Of een geveer van feedback hoeft met het kleuren van mensen niet langer precies te zijn in het benoemen van gedrag en welk effect dat heeft op de onderlinge verhouding. En dan zijn er nog verschillende kleurentesten. Hierdoor kan onduidelijkheid ontstaan omdat kleuren en kleurcombinaties in de verschillende testen vaak andere betekenissen hebben. Veel mensen komen meer dan een kleursysteem tegen en je moet oppassen dat je het nog wel over hetzelfde 'geel' hebt.

Ook kan het ontwikkelen van mensen een flinke uitdaging worden wanneer de ontvanger van feedback de eigen kleur als excuus gebruikt om niet te hoeven veranderen. De kans bestaat dat ze zich daarmee gaan vereenzelvigen waardoor het een *selffulfilling prophecy* wordt. Ze verstoppen zich achter hun resultaten: "Ik kan niet anders, zo ben ik nu eenmaal...", en: "Ze weten nu hoe ik ben." Ze gaan leven met het idee dat ze een bepaald type zijn en in het ergste geval lijden ze daaronder.

Op deze manier dragen kleuren bij aan het labelen van mensen. Dit kan resulteren in stereotyperingen en zelfs uitsluiting van mensen met een afwijkende kleur. Een ander extremitet van het gebruik van oppervlakkige testen is wanneer organisaties er zware beslissingen op baseren, bijvoorbeeld om mensen te ontslaan.

Ten slotte zijn er ook HR-professionals die zeggen dat het weinig uitmaakt wat de resultaten van zo'n kleurentest zijn; het is namelijk sowieso input voor een gesprek. Maar door een testuitslag als uitgangspunt in een gesprek te nemen, creëer je een referentiepunt voor dat gesprek. Veel loopbaanbeslissingen, het al dan niet toegewezen krijgen van functies of rollen worden hierdoor onbewust en onterecht beïnvloed. De andere kant op gebeurt ook: ik zie mensen geen interesse hebben in bepaalde functies omwille van hun kleurentestuitslagen.

Om een gesprek effectief te laten zijn is het van belang dat iedereen de betekenis van de kleuren kent. Net zoals het inzicht dat een persoon niet per definitie zijn kleur is.

### **Betere keuzes**

Vaak ingezette vragenlijsten als DISC, Insights Discovery of MBTI missen voorspellende waarde. Een tegenargument is dat de test primair de functie van spiegel moet hebben en niet voorspellend hoeft te zijn. Een spiegel moet echter ook een betrouwbaar beeld teruggeven. Anders is het, in de woorden van psycholoog Max Wildschut, een lachspiegel.

Wil iemand zich losjes oriënteren op welke testen er überhaupt zijn en globaal

kennismaken met testmateriaal? Dan zijn websites als 123test.nl of leren.nl prima. Maar zeker niet meer dan dat. Als iemand een assessment wil oefenen, dan is veruit de beste tip een goede nachtrust en fris aan het assessment beginnen.

### **Goede vragenlijsten**

Kortom, wanneer je persoonlijkheidstesten gebruikt als ondersteuning voor advies of beslissingen, zorg dan dat je goed en betrouwbaar gereedschap hebt. Maak om mensen te selecteren en ontwikkelen uitsluitend gebruik van vragenlijsten die valide en betrouwbaar zijn. Die mogen dus ook gewoon een beetje saaie uitstraling hebben en wat detectiveachtig zijn: als je ze invult weet je niet precies hoe en wat er gemeten wordt.

Hieronder vallen gevalideerde Big Five vragenlijsten die inzicht bieden in de mate van extraversie, openheid, consciëntieusheid, stabiliteit en vriendelijkheid. De HEXACO-vragenlijst

voegt hier nog een bruikbare eigenschap aan toe: eerlijkheid/nederigheid. Gebruik deze lijsten dus om persoonlijkheid in kaart te brengen. Voor slechts een paar tientjes biedt de Canadese psycholoog Jordan Peterson een uitstekende Big Five vragenlijst aan op [understandingmyself.com](http://understandingmyself.com). Ook de vragenlijsten van Hogan Assessments zijn een prima manier om de verschillende kanten van mensen effectief in kaart te brengen. Een vragenlijst als de AEM (*Attachment, Extrapolation, Maturity*) cube van Human Insight biedt uitstekend zicht op motieven, rollen, functies en de ontwikkelfase van een team of organisatie. Deze vragenlijst is overigens geen persoonlijkheidsmeting maar een drijfverentest.

Als je twijfelt over de kwaliteit van een test, raadpleeg dan de COTAN (Commissie Testaangelegenheden Nederland). Deze Commissie toetst testen op betrouwbaarheid, wetenschappelijke onderbouwing en validiteit.



### **Investeer in kennis**

Dit kan betekenen je eigen kennis verdiepen of kennis inhuren. Persoonlijkheid kan niet gereduceerd worden tot een paar kreten en kleurtjes. De gebruiker van de test mag meer inzicht krijgen in de achtergronden van een test en snappen wat de test inhoudt of zich nader verdiepen in de ins en outs van persoonlijkheid.

De testontwikkelaar mag op zijn beurt meer zijn best doen om de test op een manier te presenteren die toegankelijker is. Bijvoorbeeld door de taal meer op de gebruiker aan te passen. Spreek bijvoorbeeld niet over 'consciëntieusheid', maar over georganiseerd zijn, zelfbeheersing en detail- en doelgerichtheid.

Neem uitslagen van persoonlijkheidstesten ook nooit als ultieme waarheid. Zelfs bij gebruik van de meest gedegen testen. Check de uitslag, bijvoorbeeld in een gesprek of bij mensen die deze persoon kennen.

Denk goed na over de uitdagingen waar iemand voor komt te staan in een functie. In de praktijk wordt gedrag beïnvloed door zowel de (werk)situatie waarin iemand zicht bevindt, als door iemands persoonlijkheid. Plaats testen dus in de specifieke, relevante context. Dus welke manager krijgt iemand? Wat voor collega's of teamcultuur? In welke fase zit een afdeling of organisatie? Door kennis te hebben van wat situaties in werk van mensen vragen, kun je beter inschatten wat de relevante dimensies van het werk zijn die moeten worden meegenomen in het persoonlijkheidsprofiel.

Maak persoonlijkheidstesten onderdeel van assessment. Combineer kwalitatieve vragenlijsten met andere assessmentinstrumenten als valide en betrouwbare capaciteitentesten, simulaties en diepte-interviews. Hiermee vergroot je de voorspellende waarde van je interventie of selectie flink.

### **Tot slot**

Mensen zijn veelzijdig, wonderbaarlijk en paradoxaal. Ze vertonen verschillend gedrag in verschillende situaties en rollen. Onderdelen van mensen laten zich wel beschrijven, maar mensen vang je niet in een voorgekookte mal. Anders gezegd: zie je echt wie de mens voor je is? Door daarnaast met de persoon in kwestie in gesprek te gaan, door gedrag te observeren, verkregen informatie zorgvuldig te analyseren en uitkomsten naar relevant competentiegedrag te vertalen, zorg je ervoor dat je mensen ziet voor wie zij daadwerkelijk zijn en niet reduceert tot een testuitslag. ♦

### **Referenties**

- Boellen, J., Schröder, I., Bland, T., & Brouwer, C. (2018). *Persoonlijkheidstesten: Zin of onzin? Een kwalitatief onderzoek naar het gebruik van persoonlijkheidstesten in organisaties*. Universiteit Utrecht, [https://studenttheses.uu.nl/bitstream/handle/20.500.12932/30708/Definitief%20eindpaper%20\(1\).pdf?sequence=2](https://studenttheses.uu.nl/bitstream/handle/20.500.12932/30708/Definitief%20eindpaper%20(1).pdf?sequence=2)
- Bono, E. de (1985). *Six Thinking Hats: An essential approach to business management*. Londen: Little Brown.
- Boonstra, J. J. (2020). *Veranderen als samenspel: Een positieve kijk op het veranderen en vernieuwen van*



- organisaties. Amsterdam: Boom.
- Caluwé, L. I. A. de, & Vermaak, H. (2006). *Leren veranderen: een handboek voor de veranderkundige*. Deventer: Kluwer.
- Corr, P.J., & Matthews, G. (Eds.) (2009). *The Cambridge handbook of personality*. Cambridge: University Press.
- Engelfriet, R. (2016). *Verspil geen geld meer aan Insights Discovery, Management Drives en MBTI*. Opgehaald van <https://rengelfriet.wordpress.com/2016/11/02/neefjevanzes/>
- Engelfriet, R. (2017). *De succesillusie: Hoe trainers, goeroes en consultants u dagelijks bedriegen en hoe u daar in zeven eenvoudige stappen vanaf komt*. Zaltbommel: Haystack.
- Geisinger, K.F. (Ed) (2013). *APA Handbook of Testing and Assessment in Psychology, Vol. 3: Testing and Assessment in School Psychology and Education*. Washington: APA.
- Jung, C. G. (2003). *Psychologische typen* (11e dr.). Rotterdam: Lemniscaat.
- Lowman, R.L. (2022). *Career Assessment: Integrating Interests, Abilities, and Personality*. Washington: APA.
- Mullenders, D. (2023). Daarom doen ze dat: Kijk mee over de schouder van een psycholoog. Speciaal voor HR en leidinggevenden! Culemborg: Van Duuren.
- Reman, P., & Nordin, A. (2021). Personality tests in recruitment, [https://www.researchgate.net/publication/349466573\\_Personality\\_tests\\_in\\_recruitment](https://www.researchgate.net/publication/349466573_Personality_tests_in_recruitment)
- Vermeren, P. (2008). *De HR-ballon: 10 populaire praktijken doorprikt*. Gent: Academia.



**Danny Mullenders** is assessment- en ontwikkelingspsycholoog, personal en teamcoach. Danny studeerde cognitieve psychologie en bedrijfseconomie. Hij schreef in 2022 de bestseller *Laat je niet kiezen*. Zijn tweede boek, *Daarom doen ze dat*, ligt sinds kort in de boekhandel en is inmiddels ook een bestseller. Danny Mullenders is partner bij psychologisch adviesbureau AC&T Consulting. [www.dannymullenders.nl](http://www.dannymullenders.nl)