

Psychodiagnostiek en assessment

De praktijk van
psychologische testen, interviews
en observatiemethoden



Noordhoff Uitgevers

Henk Verhoeven

2^e druk

Psychodiagnostiek en assessment

De praktijk van psychologische testen,
interviews en observatiemethoden

Henk Verhoeven

Tweede druk

Noordhoff Uitgevers Groningen/Utrecht

Ontwerp omslag: Michiel Uilen
Omslagillustratie: antishock Adobe Stock
Hoofdstukopeningsfoto's: Photodisc

Eventuele op- en aanmerkingen over deze of andere uitgaven kunt u richten aan:
Noordhoff Uitgevers bv, Afdeling Hoger Onderwijs, Antwoordnummer 13, 9700 VB
Groningen of via het contactformulier op www.mijnnoordhoff.nl.

Aan de totstandkoming van deze uitgave is de uiterste zorg besteed. Voor informatie die desondanks onvolledig of onjuist is opgenomen, aanvaarden auteur(s), redactie en uitgever geen aansprakelijkheid. Voor eventuele verbeteringen van de opgenomen gegevens houden zij zich aanbevolen.



0 / 18

© 2018 Noordhoff Uitgevers bv Groningen/Utrecht, The Netherlands.

Behoudens de in of krachtens de Auteurswet van 1912 gestelde uitzonderingen mag niets uit deze uitgave worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen of enige andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever. Voor zover het maken van reprografische verveelvoudigingen uit deze uitgave is toegestaan op grond van artikel 16h Auteurswet 1912 dient men de daarvoor verschuldigde vergoedingen te voldoen aan Stichting Reprorecht (postbus 3060, 2130 KB Hoofddorp, www.reprorecht.nl). Voor het overnemen van gedeelte(n) uit deze uitgave in bloemlezingen, readers en andere compilatiewerken (artikel 16 Auteurswet 1912) kan men zich wenden tot Stichting PRO (Stichting Publicatie- en Reproductierechten Organisatie, postbus 3060, 2130 KB Hoofddorp, www.stichting-pro.nl).

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher.

ISBN(ebook) 978-90-01-12037-5

ISBN 978-90-01-12036-8

NUR 776

Inhoud

Inleiding 6

DEEL 1

Wat is psychodiagnostiek? 11

1 Diagnostiek: inschatten van mensen 13

- 1.1 Besliskunde 13
 - 1.2 Correlaties 16
 - 1.3 Fouten in ons beoordelingsvermogen 18
 - 1.4 Remedies tegen beoordelingsfouten 25
- Samenvatting 26

2 Kenmerken van en kwaliteitseisen aan diagnostische instrumenten 29

- 2.1 Drie soorten diagnostische instrumenten 29
 - 2.2 Wat maakt een test een goede test? 30
 - 2.3 COTAN 49
 - 2.4 Goede instrumenten vereisen goede diagnosten 50
- Samenvatting 58

DEEL 2

De praktijk: stappen in de diagnostische cyclus 61

3 Intake en probleemanalyse 63

- 3.1 Diversiteit in soorten intakegesprekken en probleemanalyses 64
 - 3.2 Luisteren: een state of mind 65
 - 3.3 Van luisteren naar analyseren 71
 - 3.4 Afbakenen van de vraagstelling 76
 - 3.5 Operationaliseren van begrippen 78
 - 3.6 Resultaat van de intake: de probleemanalyse 81
- Samenvatting 84

4 Contextkennis 87

- 4.1 Beoordelen: vergelijken met een norm 88
- 4.2 Competentieprofielen in bedrijven en organisaties 89

- 4.3 Contextkennis in ggz, onderwijs, opvoeding en zorg [95](#)
- 4.4 Professionaliteit = diagnostische kennis + contextkennis + X [99](#)
[Samenvatting 103](#)
- 5 Hypothesetoetsend model en instrumentkeuze [105](#)**
 - 5.1 Het hypothesetoetsend model (HTM) [106](#)
 - 5.2 De empirische cyclus in de wetenschap [106](#)
 - 5.3 Het hypothesetoetsend model: startpunt in de instrumentkeuze [108](#)
 - 5.4 Instrumenten [110](#)
 - 5.5 Definitieve instrumentkeuze [156](#)
[Samenvatting 160](#)
- 6 Testafname, meting en scoring [163](#)**
 - 6.1 De testdag [163](#)
 - 6.2 Keuze normtabellen [165](#)
 - 6.3 Scoren [166](#)
[Samenvatting 174](#)
- 7 Verwerking en interpretatie van onderzoeksgegevens [177](#)**
 - 7.1 Interpretatie testgegevens: van vaktaal naar lektentaal [177](#)
 - 7.2 Empirische relaties tussen simpele testscores en de complexe werkelijkheid [180](#)
 - 7.3 De betekenis van enkelvoudige en gecombineerde testuitslagen [183](#)
 - 7.4 Circumplexen en het brainstormdocument: onmisbare hulpmiddelen [186](#)
[Samenvatting 190](#)
- 8 Rapporteren en adviseren [193](#)**
 - 8.1 Het psychologische rapport [193](#)
 - 8.2 Het technisch-professionele verslag [194](#)
 - 8.3 Het cliëntverslag/cliëntrapport [195](#)
 - 8.4 Nabespreking en advisering [201](#)
[Samenvatting 206](#)
[Bijlage: voorbeeld van een cliëntrapport naar aanleiding van een developmentcenter 207](#)
- 9 Veelgebruikte testen in Nederland [213](#)**
 - 9.1 Bevoegdheid met betrekking tot testgebruik [214](#)
 - 9.2 Prestatie Motivatie Test (PMT) [214](#)
 - 9.3 Nederlandse Persoonlijkheidsvragenlijst (NPV) [216](#)
 - 9.4 Wechsler Adult Intelligence Scale, Nederlands. Versie 4 (WAIS-IV-NL) [219](#)

- 9.5 Differentiële Aanleg Test voor Human Resources Management (DAT voor HRM) [221](#)
- 9.6 Amsterdamse Beroepen Interesses Vragenlijst (ABIV) [222](#)
- 9.7 NEO Personality Inventory, versie 3 (NEO-PI-3) [223](#)
- 9.8 Symptoms Checklist (SCL-90) [225](#)
- 9.9 Competentie-Belevingsschaal voor Adolescenten (CBSA) [226](#)
- 9.10 Utrechtse Coping Lijst (UCL) [227](#)
- 9.11 Utrechtse Burnout Schaal (UBOS) [228](#)
[Samenvatting](#) [230](#)

Deel 3

Diagnostiek; politiek, geschiedenis en de keerzijde van testgebruik [233](#)

10 Diagnostiek is meer dan testen afnemen [235](#)

- 10.1 Vak- of kerncompetenties en beroepsrollen van de toegepast psycholoog [235](#)
- 10.2 Van problemen naar oplossingen [237](#)
- 10.3 Diffuse belangen en politieke processen [241](#)
[Samenvatting](#) [243](#)

11 De geschiedenis van de psychodiagnostiek [245](#)

- 11.1 Aanloop naar een systematische diagnostiek [245](#)
- 11.2 De psychologische test: kind van de industriële revolutie [246](#)
- 11.3 Pseudowetenschap en psychodiagnostiek [250](#)
[Samenvatting](#) [258](#)

12 De keerzijde van testgebruik [261](#)

- 12.1 Misbruik IQ-testen [261](#)
- 12.2 De antitestbeweging [263](#)
- 12.3 De controverser persoon versus situatie [265](#)
- 12.4 Persoon versus situatie: antwoord van de testpsychologie [268](#)
- 12.5 Sjoemelsoftware, Teaching to the Test, de lakei en de koning [269](#)
- 12.6 Overdiagnoses [274](#)
- 12.7 Identiteit [276](#)
- 12.8 Tot besluit [279](#)
[Samenvatting](#) [280](#)

Literatuur [281](#)

Register [287](#)

Over de auteur [292](#)

Inleiding

De psychologische test is overal

De psychologische test is niet meer weg te denken uit onze moderne maatschappij. Iedereen is gedurende zijn leven wel ten minste eenmaal onderworpen aan een formeel psychologisch onderzoek en velen maken het verschillende keren mee. Van Minden (1995) komt op een schatting van 515.000 mensen per jaar die getest worden (150.000 tijdens sollicitatieprocedures, 85.000 voor militaire keuringen, 92.000 voor de Cito-toets, 150.000 voor school- en beroepskeuze en een restcategorie van 38.000). Mogelijk heeft hij de getallen wat ruim ingeschat en zijn de exacte cijfers inmiddels wat veranderd (keuringen voor de dienstplicht bestaan bijvoorbeeld niet meer), maar feit blijft dat de test in alle uithoeken van onze samenleving is doorgedrongen.

Belangrijke beslissingen over mensen worden – gelukkig – steeds meer gebaseerd op de uitkomsten van objectief onderzoek met gedegen professionele instrumenten. Een belangrijk deel van de willekeur van weleer is daarmee uitgebannen (Dijkman, 1998), want hoe je het ook wendt of keert, er moeten beslissingen over mensen genomen worden (door henzelf en/of door anderen) en dit gebeurt altijd op basis van (gedegen of minder gedegen) inschattingen van eigenschappen en kwaliteiten. Dus: waar jij als professional kunt bijdragen aan de kwaliteit van deze beslissingen, moet je dat niet nalaten.

Toepassingen van psychodiagnostisch onderzoek

Waar komen we praktische toepassingen van testen zoal tegen? En waar ontmoeten we de toegepast psycholoog en andere professionals die menselijke eigenschappen en talenten inschatten in hun rol van psychodiagnost? Enkele voorbeelden uit verschillende werkvelden (Asendorpf, 2011; Smit, Verhoeven & Driessen, 2006; Ter Laak & De Goede, 2003):

- Selectie van personeel.
- Advisering aan individuen over de best passende functies of loopbanen.
- Ondersteunend aan outplacement en re-integratieprocessen, management development.
- Effectmetingen opleidingsactiviteiten.
- Samenstellen van (management)teams.
- Preventie in de gezondheidszorg. Bijvoorbeeld: op basis van persoonlijkheidsonderzoek kan vastgesteld worden wie gevoelig is om verslaafd te raken, een hartaanval te krijgen of zwaarlijvigheid te ontwikkelen. Aan deze mensen wordt dan een preventief programma aangeboden.
- Marketeers gebruiken psychologische testen om persoonlijkheidskenmerken, voorkeuren, koopgedrag en affecten van specifieke doelgroepen te onderzoeken, zodat productontwikkeling en reclame daarop beter kunnen worden afgestemd.

- Schoolkeuze en plaatsing in speciaal onderwijs (de Cito-toets heeft sinds zijn ontstaan veel bijgedragen aan een juiste schoolkeuze van basisschoolleerlingen).
- Diagnose van leer- en studieproblemen (dyslexie, faalangst, ADHD).
- Beslissingen over uithuisplaatsing van verwaarloosde, misbruikte of mishandelde kinderen.
- Matchen van patiënten aan specifieke therapievormen in de ggz (individueel of groepsgewijs, meer cognitief of juist gedragsgericht, intra- of extramuraal et cetera).
- Diagnostiek van psychiatrische en neurologische aandoeningen.
- Vaststellen van de oorzaak van leer- of gedragsproblemen (dyslexie, autisme, ADHD, ADD en dergelijke).
- Matching van koppels of huwelijkspartners.
- Diagnostiek ten behoeve van expertiserapporten in strafzaken (de bekende uitspraken over of een verdachte al dan niet toerekeningsvatbaar was toen hij de misdaad beging). Ondersteunend onderzoek ten behoeve van de vraag of een veroordeelde tbs moet krijgen of bijvoorbeeld op vervroegd verlof mag.
- Daderprofilering. Hierbij wordt de psychodiagnostiek gebruikt om een beeld te krijgen van (de nog onbekende) dader van een misdrijf. Hoewel in de populaire media en in Amerikaanse politieseries op tv de mogelijkheden van daderprofilering nogal eens worden overdreven, draagt deze benadering wel degelijk bij aan de opsporing van misdadigers.

Op basis van interviews, psychologische testen en observatietechnieken – de drie methoden die we samen aanduiden met de term psychodiagnostiek – worden dus tal van beslissingen genomen. Ingrijpende en zwaarwegende beslissingen, maar ook beslissingen met minder impact en minder verstreckende consequenties.

Opbouw van dit boek

Dit studieboek is geschreven voor toegepast psychologen, hrm-adviseurs en andere professionals die beroepsmatig persoonlijkheid, kwaliteiten, vaardigheden en ontwikkelpunten van mensen in kaart brengen. Vaak zijn dit studenten die een hbo-opleiding of universitaire studie volgen en na hun bachelor of master in de praktijk aan de slag gaan met concrete vraagstukken rondom menselijk gedrag en daarbij gebruikmaken van psychodiagnostische competenties (naast competenties als trainen, coachen, voorlichten, adviseren, onderzoeken et cetera). Het vaststellen van psychiatrische ziektebeelden of neurologische aandoeningen zal bijvoorbeeld de toegepast psycholoog niet zelfstandig doen, dit valt buiten zijn bevoegdheden. Wel kan hij hierin een ondersteunende rol vervullen. Maar er zijn echter tal van diagnostische vraagstukken die hij wel zelfstandig aanpakt.

Dit boek beoogt de student in de toegepaste psychologie en andere vakgebieden een inleiding te bieden in het boeiende vak psychodiagnostiek. Het boek is opgebouwd uit drie delen. In het eerste deel wordt ingegaan op menselijke beslissingsprocessen en op welke manier die terugkomen in de professionele psychodiagnostiek. Verder komen de kenmerken van verschillende instrumenten aan de orde en beschrijven we de kwaliteitscriteria die hieraan worden gesteld. In deel twee volgen we het dagelijkse werk van de psychodiagnost. We bespreken chronologisch welke stappen hij doorloopt om uiteindelijk het probleem of de vraag van de cliënt of opdrachtgever helder in kaart te brengen. In het laatste hoofdstuk van dit deel beschrijven we

tien veelgebruikte testen in Nederland. In deel drie tot slot, gaan we in op de geschiedenis van ons vak. Doorgaans is dit waar studieboeken mee beginnen. Wij hebben ervoor gekozen om dit anders aan te pakken; je kunt deze onderwerpen namelijk pas goed begrijpen wanneer je inzicht verworven hebt in de praktische gang van zaken. Het boek eindigt met een waarschuwing. Psychodiagnostiek is een nuttig en maatschappelijk belangrijk vak. Je kunt er veel goeds mee doen. Maar het kan ook fout gaan. Daarover gaat het laatste hoofdstuk, als waarschuwing om steeds alert te blijven op de implicaties van je handelen.

Bij dit boek hoort een website: **www.psychodiagnostiek.noordhoff.nl**.

Dagelijks worden in onze samenleving beslissingen over mensen genomen. Vaak door henzelf, maar soms ook door anderen. Denk aan het schoolhoofd die beslist naar welk type voortgezet onderwijs een kind gaat, de personeelsfunctionaris die bepaalt welke sollicitant de felbegeerde functie in de wacht sleept of de toegepast psycholoog die beslist welke coaching of training zijn cliënt het meest vooruit zal helpen. Psychodiagnostiek gaat over het nemen van dergelijke beslissingen met speciaal daarvoor ontwikkelde diagnostische instrumenten. De grens van ons vak is dus niet altijd duidelijk. Doet alleen de psycholoog aan psychodiagnostiek? Of ook de personeelsfunctionaris en het schoolhoofd? Het antwoord is beide.

Dit boek is geschreven voor iedereen die op professionele wijze eigenschappen van mensen moet inschatten, dus voor toegepast psychologen, hrm-deskundigen, pedagogen en maatschappelijk werkers. Voor iedereen geldt dat professionele diagnostiek altijd kan helpen om de kwaliteit van de uiteindelijke beslissingen naar een hoger niveau te tillen.

Professionele psychodiagnostiek lijkt sterk op wat we dagelijks doen wanneer we mensen inschatten en op basis daarvan beslissingen nemen. Toch zijn er een aantal belangrijke verschillen. Om die verschillen duidelijk te maken zullen we in deel 1 van dit boek ingaan op de wijze waarop in het algemeen menselijke beslissingen tot stand komen en op welke punten dit kan mislopen. Daarna behandelen we de instrumenten die de psychodiagnost inzet om deze fouten te voorkomen en onderzoeken we aan welke kwaliteitseisen die instrumenten moeten voldoen. Daarbij besteden we uiteraard ook aandacht aan de diagnost zelf, want zonder zijn of haar vakmanschap is zelfs het beste instrument een nutteloos ding. Langzaam zal duidelijk worden waarin beslissingsprocessen van de leek verschillen van die van de professionele toegepast psycholoog.

DEEL 1

Wat is psychodiagnostiek?

- 1 Diagnostiek: inschatten van mensen 13
- 2 Kenmerken van en kwaliteitseisen aan diagnostische instrumenten 29



1

Diagnostiek: inschatten van mensen

- 1.1 Besliskunde
- 1.2 Correlaties
- 1.3 Fouten in ons beoordelingsvermogen
- 1.4 Remedies tegen beoordelingsfouten

Diagnostiek gaat over het inschatten van mensen: wat zijn hun persoonlijkheidskenmerken, waarden en normen, vaardigheden, attitudes en typische manieren van doen? Dagelijks maken wij allemaal inschattingen van situaties en mensen om ons heen. Soms doen we dat bewust, soms onbewust. We hebben deze inschattingen nodig omdat we op basis hiervan beslissingen nemen. Wanneer je inschat dat je bij een bepaalde docent maar beter niet onvoorbereid in de les kunt verschijnen en dat dit bij een andere docent waarschijnlijk geen probleem oplevert, dan doe je eigenlijk al aan psychodiagnostiek. Je schat immers persoonskenmerken en toekomstig gedrag van die docenten in.

Theorieën omtrent besliskunde kunnen ons veel leren over de manier waarop beoordelingen en beslissingen tot stand komen. Omdat wij als mensen niet altijd automatisch de juiste beslissingen nemen, is het goed om wat dieper in te zoomen op de meest voorkomende fouten en valkuilen in ons beoordelingsvermogen. Verder zullen we in dit hoofdstuk kijken naar een aantal remedies tegen onjuiste oordeelsvorming en concluderen dat psychodiagnostiek een gestructureerde aanpak is om het proces van oordeelsvorming over menselijke kwaliteiten en gedragingen in juiste banen te leiden en de belangrijkste fouten hierin te voorkomen.

Valkuilen in ons
beoordelings-
vermogen

1.1 Besliskunde

Niet alleen degenen die toegepaste of academische psychologie hebben gestudeerd, zijn psycholoog. Op een bepaalde manier zijn alle mensen psycholoog. Iedereen heeft ideeën over hoe andere mensen in elkaar steken, wat ze waarom doen, hoe ze op specifieke situaties zullen reageren en hoe hun karakter is.

De leek als
psycholoog

De meeste mensen zijn gelukkig ook vrij goede psychologen, want het lukt ons in het dagelijks leven doorgaans vrij aardig te anticiperen op wat anderen gaan doen. Vooraf kunnen voorspellen wat er zal gaan gebeuren, dat is altijd het beste bewijs voor het feit dat je werkelijk begrijpt hoe iets in elkaar steekt. Dit geldt zowel voor de werking van een automotor, als voor de persoonlijkheid en de situationele factoren die het doen en laten van mensen bepalen.

Psychologen zijn niet de enigen die voortdurend voorspellingen moeten doen. Denk eens aan de volgende mensen, beroepen en rollen:

- schoolhoofd
- arts
- weerman
- beleggingsadviseur
- verzekeringsadviseur
- marketeer
- personeelsfunctionaris
- manager
- rechter
- echtgenoot
- koper van een auto

Voorspellingen

Probeer zonder vooruit te lezen, te bedenken wat voor soort voorspellingen zij allemaal doen. Is het je gelukt?

Het schoolhoofd bekijkt of Jantje naar de havo, het vwo of het vmbo moet. Hij doet dus een voorspelling, bijvoorbeeld: Jantje zal succesvol zijn op het vwo. De arts die de diagnose kanker stelt, doet de voorspelling dat wanneer een chirurg zijn patiënt opereert, die ook werkelijk een kankergezweel zal aantreffen. De weerman voorspelt het weer van morgen. De beleggingsadviseur doet op basis van zijn kennis van een bedrijfstak een voorspelling over de vraag of een bedrijf over een jaar nog winstgevend zal zijn en beveelt op basis hiervan zijn cliënt aan aandelen van dit bedrijf te kopen of verkopen. De verzekeringsadviseur schat de kans in dat een huis zal afbranden. De marketeer kiest een product en doelgroep, en voorspelt dat dit tot hoge verkoopcijfers zal leiden. Na enkele jaren weet hij of zijn voorspelling juist was of niet. De personeelsfunctionaris voorspelt op basis van een selectiegesprek of een sollicitant een goede medewerker zal zijn, of zijn afdelingsmanager tevreden met hem zal zijn, of hij zijn omzet haalt en er niet met de kas vandoor zal gaan. Een rechter geeft een tbs'er verlof, omdat hij voorspelt dat de gedetineerde niet de fout in zal gaan. Soms klopt zijn voorspelling niet en pleegt de veroordeelde misdadiger tijdens zijn verlof nogmaals een moord. Ook wanneer je trouwt of gaat samenwonen, doe je een voorspelling, namelijk dat jij en je partner na tien, twintig of vijftig jaar nog steeds vinden dat je gelukkig met elkaar zult zijn.

Test en criterium

Merk op dat in elk voorspellingsproces steeds twee variabelen een hoofdrol spelen, namelijk een beoordelingsmoment nu en een beoordelingsmoment ergens in de toekomst. In de psychodiagnostiek noemen we dit test en criterium. De mate waarin deze twee beoordelingen met elkaar overeenstemmen, zegt iets over de kwaliteit van je voorspellingen en je beoordelingsvermogen. Een weerman die steeds moet concluderen dat het regent wanneer

hij zonneshijn voorspelde of windstilte constateert wanneer hij een storm had aangekondigd, verliest snel aan geloofwaardigheid. Bij eenmalige voorspellingen – bijvoorbeeld wanneer je met iemand gaat trouwen – is het lastiger, maar wanneer je bepaalde voorspellingen vaak doet, is het mogelijk aan de hand van de juistheid van je voorspellingen in het verleden een uitspraak te doen over jouw bekwaamheid in de vaardigheid van het voorspellen. We doen dan een voorspelling over hoe goed jij als voorspeller zult functioneren. We noemen dit ‘predictieve validiteit’. Ook van een psychologische test kunnen we de predictieve validiteit vaststellen (zie subparagraaf 2.2.4).

Predictieve
validiteit

Op basis van de twee beoordelingsmomenten die we eerder bespraken, worden er in de besliskunde vier soorten beslissingen onderscheiden. Neem als voorbeeld de weerman. Hij doet een voorspelling over de vraag of het morgen wel of niet zal gaan regenen. De vraag ‘Regent het morgen?’, beantwoordt hij met ‘ja’ of ‘nee’. De volgende dag kunnen we werkelijk vaststellen of het regent, dus ook hier kan een ‘ja’- of ‘nee’-antwoord volgen. Wanneer we het voorspellingsmoment (vandaag) – het testmoment of wel test genoemd – en het toetsmoment (morgen) – het criterium – in een kwadrant zetten, dan leidt dat tot een besliskundematrix zoals is weergegeven in figuur 1.1.

FIGUUR 1.1 Besliskundematrix

Dag na de voorspelling: regent het?	Ja	FN	VP
	Nee	VN	FP
		Nee	Ja
		Voorspelling: gaat het morgen regenen?	

We beginnen rechtsboven in het kwadrant. Hier heeft de weerman voorspeld dat het gaat regenen, en het is ook werkelijk gaan regenen. Dit noemen we een ‘Valid Positive’ (VP). De voorspelling was positief (*Positive*) en bleek later ook waar (*Valid*). De beslissingen rechtsonder noemen we een ‘False Positive’ (FP). De voorspelling was positief (*Positive*), maar bleek niet te kloppen (*False*). Ook linksboven staat een foute beslissing. De weerman voorspelde dat het niet zou gaan regenen (*Negative*), maar dit bleek niet te kloppen (*False*), het ging namelijk wel regenen. Dit is dus een ‘False Negative’ (FN). Tot slot staat linksonder de ‘Valid Negative’ (VN), de voorspelling was dat het niet zou gaan regenen en het bleef ook werkelijk droog.

Valid Positive
False Positive

False Negative
Valid Negative

Duidelijk zal zijn dat de kwaliteit van de voorspeller toeneemt naarmate hij meer Valids scoort en zo min mogelijk False beslissingen neemt. Een voorspeller kan een persoon zijn, maar ook een test, of een hele voorspellingsprocedure, zoals een testbatterij of een assessmentprogramma. We kunnen de kwaliteit van het voorspellend vermogen van een instrument of een voorspeller dus uitdrukken in een getal. Bijvoorbeeld, je kunt concluderen dat weerman A deze maand 80% valide voorspellingen heeft gedaan en 20% verkeerde (False) voorspellingen deed en dus beter is dan weerman B die maar 60% juiste voorspellingen deed en er in 40% van de gevallen naast zat.

1.2 Correlaties

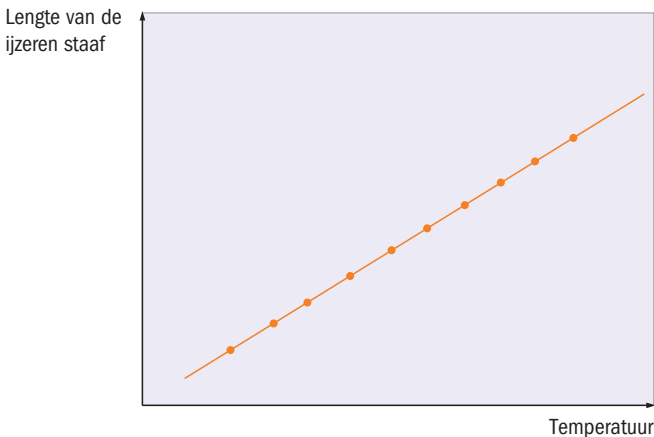
Zolang voorspellingen van de soort ja/nee zijn, is het percentage zoals aan het eind van de vorige paragraaf is beschreven een goede maat om de voorspellende (predictieve) kracht van een weerman uit te drukken. Lastiger wordt het wanneer we het vermogen van de weerman om de temperatuur te voorspellen willen vaststellen. Wanneer de weerman 23 °C voorspelde en het wordt 24 °C, zit hij er in principe naast, maar dit is natuurlijk minder erg dan dat het de volgende dag 33 °C of juist 10 °C bleek te zijn. We hebben dus een fijnere maat nodig om de voorspellende kwaliteiten van weermanen op dit punt vast te stellen.

Correlatiecoëfficiënt

Statistici hebben voor dergelijke problemen het begrip correlatiecoëfficiënt ontwikkeld. Dit begrip speelt een niet weg te denken rol in de psychodiagnostiek. De correlatiecoëfficiënt is een getal tussen de 0 en 1.00 dat de sterkte van de relatie tussen twee variabelen aangeeft.

In de natuurkunde treffen we perfecte relaties aan. Een ijzeren staaf zet uit wanneer we die warmer maken. Wanneer we honderd metingen verrichten en daarbij steeds de temperatuur noteren en de lengte van de ijzeren staaf, dan kunnen we die in een grafiek uitzetten. Op de x-as staat de temperatuur, op de y-as de lengte van de ijzeren staaf (zie figuur 1.2).

FIGUUR 1.2 Relatie tussen de temperatuur en de uitzetting van een ijzeren staaf

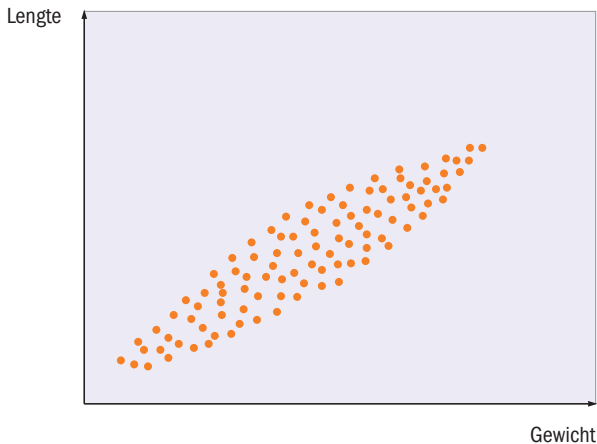


In figuur 1.2 zien we dat alle stippen op één strakke lijn liggen. Wanneer iemand ons vraagt hoe lang de staaf precies is bij een temperatuur die we niet gemeten hebben, dan kunnen we uit de grafiek afleiden hoe lang de staaf is. Het blijkt dus mogelijk een variabele te voorspellen uit een andere variabele. Omdat alle meetpunten op exact dezelfde rechte lijn liggen, zeggen we dat de correlatiecoëfficiënt tussen X en Y (tussen warmte en lengte van de staaf dus) 1.00 is. Een perfecte correlatie.

Niet-perfecte correlaties

In de menswetenschappen zullen we nooit dergelijke perfecte relaties aantreffen. We weten bijvoorbeeld dat er een verband bestaat tussen lichaamslengte en gewicht. Langere mensen zullen over het algemeen zwaarder zijn dan kleinere mensen. Maar er bestaan natuurlijk ook uitzonderingen op deze regel. Kleine, dikke mensen kunnen zwaarder zijn dan langere, dunne mensen. Toch neemt dit niet weg dat we in het algemeen kunnen stellen dat langere mensen ook zwaarder zijn. Hoe sterk de relatie tussen deze twee variabelen is, kan worden uitgedrukt in een correlatiecoëfficiënt. Onderzoekers hebben daartoe van duizend mensen de lengte gemeten en het gewicht bepaald. Als we al deze metingen in een grafiek zetten krijgen we een soort puntenwolk (zie figuur 1.3).

FIGUUR 1.3 Lengte- en gewichtsgegevens van duizend mensen

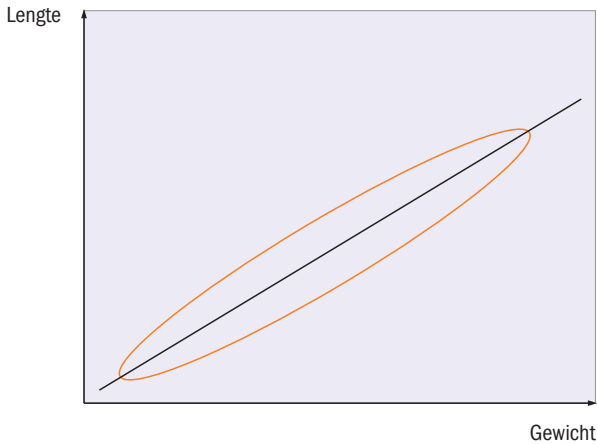


Duidelijk zal meteen zijn dat het niet een willekeurige puntenwolk is. Er zit een patroon in. Met behulp van statistische technieken kunnen we een lijn tekenen die precies midden door de puntenwolk gaat (dat wil zeggen dat de opgetelde afstanden van alle punten tot die lijn, zo klein mogelijk gemaakt is). Ook kunnen we een soort ellips tekenen die aangeeft hoe plat of juist bol de verzameling punten is (zie figuur 1.4).

Puntenwolk

Deze platheid of bolheid kunnen we in een getal uitdrukken. Dit getal is de correlatiecoëfficiënt waar we net over spraken. Nadert de coëfficiënt het getal 1.00, dan is de ellips zeer plat. Totdat uiteindelijk een strakke lijn ontstaat wanneer de correlatiecoëfficiënt 1.00 is. Hoe boller de ellips hoe lager de correlatiecoëfficiënt. Wanneer die 0.00 is, is er totaal geen relatie

FIGUUR 1.4 Ellips die visueel de puntenwolk van lengte- en gewichtsgegevens weerspiegelt



tussen twee variabelen. Wanneer die 0.5 is, is er een behoorlijke relatie, maar kunnen we slechts schattingen maken van één variabele wanneer de andere bekend is.

De correlatie tussen lengte en gewicht is overigens 0.80, dus vrij hoog. Wanneer we het gewicht van iemand weten, kunnen we niet exact voorspellen wat zijn lengte is, maar we kunnen wel een vrij nauwkeurige schatting daarvan maken.

Duidelijk zal zijn dat de correlatiecoëfficiënt in de wetenschap een zeer belangrijke getal is. We kunnen dan namelijk variabelen inschatten die in de toekomst gaan plaatsvinden of variabelen die niet direct zichtbaar zijn. Wanneer we bijvoorbeeld weten dat de correlatiecoëfficiënt tussen een IQ-test en succes op het werk als manager 0.55 bedraagt, dan weten we dat iemand die hoog scoort op een IQ-test waarschijnlijk een betere manager zal worden dan iemand die laag scoort. Omdat de correlatie 0.55 is, zullen ook andere factoren meespelen in de kans dat iemand succesvol wordt. Wanneer we die andere factoren ook weten en weten hoe hoog de correlatie daarvan is met werksucces, dan kunnen we de laatste variabele steeds nauwkeuriger inschatten en dus steeds betere voorspellingen doen.

De psychodiagnostiek is voortdurend op zoek naar persoonlijkheidsvariabelen waarmee andere variabelen voorspeld kunnen worden. Bijvoorbeeld: succes als manager, de kans dat een leerling succesvol de havo afmaakt, de kans dat een echtpaar zal gaan scheiden, de kans dat iemand aan alcohol verslaafd raakt, of zelfmoord pleegt, of de kans dat een specifieke therapie aanslaat bij een depressieve vrouw.

Inschatten toekomstige variabelen

Speuren naar voorspellende variabelen

1.3 Fouten in ons beoordelingsvermogen

Dat we in het dagelijkse leven best goede beslissingen nemen, zowel over mensen als over tal van andere zaken, maakt dat we doorgaans vrij ongeschonden de dag doorkomen. Een tochtje door het verkeer vraagt al vele tientallen, zo niet honderden beslissingen die ons het leven zouden kunnen

kosten als we er ernstige fouten bij zouden maken. De mens is doorgaans geen slechte diagnost. Dat geldt vooral voor praktische zaken en veelvoorkomende aspecten (waarmee we dus veel hebben kunnen oefenen). Maar dat ligt dat toch wat anders wanneer we meer ingewikkelde problemen moeten inschatten, of uitspraken moeten doen over onderwerpen waarmee we maar weinig te maken krijgen of gebeurtenissen die verder weg in de toekomst zullen plaatsvinden.

De sociale psychologie heeft ons gewezen op structurele mankementen in de manier waarop we tot inschattingen en voorspellingen komen. De toegepaste psycholoog als diagnost moet van deze fouten op de hoogte zijn en er extra alert op zijn ze te voorkomen. Verder moet hij de mechanismen koesteren die in zijn instrumentarium zijn ingebouwd om hem bij deze lastige opgave te ondersteunen. In geen geval mag hij deze ingebouwde veiligheidskleppen moedwillig uitzetten!

Hierna zullen we een aantal van de meest voorkomende fouten bespreken, te weten:

- 1 verstandige fouten
- 2 overschatten van specifieke kansen
- 3 beschikbaarheidsheuristiek
- 4 regressie naar het gemiddelde
- 5 eerste en laatste indruk
- 6 voorbarige reductie van cognitieve dissonantie

1.3.1 Verstandige fouten

Een aantal van de beoordelingsfouten die we maken, zouden we goede of verstandige fouten kunnen noemen. Ze helpen ons problemen te voorkomen. Neem een hert, ook dat dier moet beslissingen nemen. Wanneer het ergens rustig staat te grazen en het hoort geritsel in het struikgewas, dan moet het inschatten – dus voorspellen – of dat geritsel van de wind afkomstig is of dat er een leeuw op hem loert. Als er echt een leeuw zit en het hert vlucht weg, of het blijft staan en er zit geen leeuw, dan heeft het valide beslissingen genomen. Voorspelt het dat er wel een leeuw tevoorschijn zal komen, terwijl dat niet zo is, dan vlucht het voor niks. Zijn beoordeling was dan een False Positive. Het zou ook een False Negative (FN) kunnen maken: het denkt dat er geen leeuw zit, maar er zit er wél een.

Duidelijk zal zijn dat een FN dodelijk is voor een hert, terwijl een FP alleen een nodeloze renpartij kost. Een hert maakt dan ook liever duizend FP's dan één FN. Niet alle fouten zijn dus even kostbaar. In vaktaal: een hert heeft de aftestgrens – de verticale middenlijn in figuur 1.1, waarvan links de nee-voorspelling en rechts de ja-voorspelling staat – extreem naar links verschoven. Het hert zal het minste geritsel interpreteren als de aanwezigheid van een leeuw, om zo het aantal FN's tot een minimum te beperken. Het is echter altijd zo dat vermindering van het aantal FN's ook het aantal FP's, het aantal nodeloze renpartijen, vergroot.

Elke beslissing heeft twee kanten

In de mensenwereld is dat net zo. Wanneer er een gruwelijke moord gepleegd is, roept het publiek dat rechters strenger moeten straffen en verdachten minder makkelijk vrijgesproken mogen worden. Het grote publiek wil dat iemand die schuldig is, ook echt wordt gestraft en niet door een vormfout of gebrek aan bewijs op vrije voeten komt.

Wat men daarbij soms over het hoofd ziet, is dat met het verminderen van het aantal mensen dat schuldig is maar toch wordt vrijgesproken (FN's), het

Ethische discussie

aantal mensen dat onschuldig is en toch wordt veroordeeld (de FP's) evenredig toeneemt. Bij een discussie over de vraag of we de afteftgrens naar links moeten verschuiven, dienen beide beoordelingsfouten in ogenschouw genomen te worden. Wat vinden we erger? Iemand die schuldig is op vrije voeten? Of iemand die onschuldig is, in de gevangenis? Dit is een ethische discussie die in het maatschappelijk debat gevoerd moet worden, de beslistkundige (in casu een toegepast psycholoog) kan alleen maar aangeven wat de consequenties zijn van bepaalde beslissingen, net zoals hij dat doet in andere situaties waarin afteftgrenzen worden gekozen.

Een voorbeeld van een situatie waarbij de afteftgrens extreem naar rechts wordt verschoven is de selectie van toekomstige straaljagerpiloten. Het aantal jongeren dat op een dergelijke vacature solliciteert is zeer groot, terwijl er maar enkele plekken per jaar beschikbaar zijn en de kosten van een opleiding zeer hoog zijn. Afvallers tijdens de opleiding kosten defensie handenvol geld. Dus, wat doet de luchtmacht? Zeer streng selecteren. Om aangenomen te worden doorlopen rekruten een ongewoon streng en uitgebreid testprogramma. Alleen de allerbesten worden doorgelaten (dat wil nu nog zeggen de rekruten die het in het testprogramma het best doen). De luchtmacht probeert het aantal FP's tot een minimum te reduceren, omdat een afgebroken opleiding erg veel geld kost. Gevolg hiervan is wel dat het aantal FN's ook zeer groot is. De laatste groep bestaat dus uit jongens en meisjes die volgens de test te licht bevonden zijn, maar feitelijk wel geschikt waren om straaljagerpiloot te worden. Kortom, het maken van fouten kan soms een bewuste en verstandige keuze zijn (afhankelijk natuurlijk van uit wiens positie je het bekijkt). Realiseer je dat dergelijke afwegingen van de kosten van foute beslissingen overal in je leven voorkomen: bij de keuze van een opleiding, een baan, een levenspartner of heel eenvoudig of je nog net door het oranje of rode stoplicht kunt rijden of niet.

Kosten van foute beslissingen**1.3.2 Overschatting van specifieke kansen**

De kans dat bepaalde gebeurtenissen zich voordoen of de kans dat iemand over een bepaalde eigenschap beschikt, wordt soms verkeerd ingeschat. De psychologen en Nobelprijswinnaars Amos Tversky en Daniel Kahneman (Kahneman, 2011) legden aan hun proefpersonen de volgende beschrijving voor:

Linda is 31 jaar oud, vrijgezel, erg uitgesproken en slim. Ze is afgestudeerd psychologe. Als student was ze erg betrokken bij onderwerpen als discriminatie en sociale rechtvaardigheid en ze nam deel aan antinucleaire demonstraties.

Vervolgens vroegen Tversky en Kahneman aan hun proefpersonen welke van de volgende uitspraken het meest waarschijnlijk is:

- 1 Linda werkt bij een bank.
- 2 Linda werkt bij een bank en is actief in de vrouwenbeweging.

Wat zou jij antwoorden? Hetzelfde als wat 85% van de proefpersonen deden en optie 2 kiezen? Of zou je voor optie 1 gaan? Laten we je uit je onzekerheid halen. Optie 1 is het meest waarschijnlijk. Niet omdat het korte verhaal hierboven erg veel over Linda zegt, en zeker te weinig is om iets over haar beroeps- en politieke leven te kunnen zeggen, maar simpelweg omdat vanwege een basale statistische wet de kans op twee samenvallende feiten altijd kleiner is dan de kans op één daarvan. Een bruine hond is

ook altijd een hond, maar niet elke hond is bruin; de kans dat de familie Jansen als huisdier een hond heeft, is altijd groter dan de kans dat hun huisdier een hond is én dat die hond bruin is. De kans om eenmaal munt te gooien is altijd groter dan de kans op eenmaal munt én eenmaal kop. Hoe vanzelfsprekend dit ook lijkt, we maken steeds weer de fout die ook 85% van Tversky en Kahnemans proefpersonen maakten.

Een tweede experiment van beide onderzoekers begon ook weer met een verhaaltje:

Jan is verlegen en teruggetrokken, toont weinig interesse in anderen, is zeer intelligent, punctueel en georganiseerd, en hij leidt een regelmatig leven.

Wat is het meest waarschijnlijke beroep dat Jan uitoefent?

- 1 schrijver van wiskundeboeken
- 2 piloot
- 3 vertegenwoordiger
- 4 boer
- 5 beroepsvoetballer

Veel proefpersonen antwoorden met schrijver van wiskundeboeken. Waarom? Omdat dat het meest bij het karakter van Jan lijkt te passen. Deze respondenten vergeten echter dat er maar zeer weinig schrijvers van wiskundeboeken zijn en veel meer boeren of vertegenwoordigers. De kans dat dit het beroep van Jan is, is vele malen groter dan dat hij schrijver is (en hiervan weten we dat die kans groter is dan dat hij schrijver van wiskundeboeken is). De beoordelingsfout die hier gemaakt wordt, is dat veel mensen afgaan op wat intuïtief logisch klinkt (of zouden we moeten zeggen 'psycho-logisch' klinkt?) en een heel belangrijk gegeven vergeten, namelijk de *base rate*. Wat dit belangrijke begrip precies inhoudt, komt in subparagraaf 2.4.2 aan de orde.

Base rate

1.3.3 Beschikbaarheidsheuristiek

Onderzoekers ontdekten dat we gebeurtenissen veel zwaarder laten wegen dan dingen die niet gebeurd zijn, zelfs als die juist daarom des te meer bijzonder zijn. Wanneer *De Telegraaf* groot kopt dat er ergens een moord gepleegd is, is heel Nederland geschokt. Veel mensen schudden hun hoofd en vragen zich af waar het toch met ons land naartoe moet of durven 's avonds niet meer alleen over straat. Wat deze mensen over het hoofd zien is het feit dat in Nederland moordcijfers historisch laag zijn, überhaupt dat de kans dat we in de westerse wereld door geweld omkomen in de menselijke geschiedenis nog nooit zo klein is geweest (Pinker, 2011). We laten ons dus bang maken door schijnbaar bijzondere, maar statistisch gezien feitelijk irrelevante informatie.

Opvallende informatie, maar daardoor niet noodzakelijk belangrijke informatie, laten we onevenredig zwaar meewegen in onze oordeelsvorming. Bij de selectie van sollicitanten weegt dat iemand topsporter is geweest of een strafblad bezit veel zwaarder dan eigenlijk verantwoord is. Dit komt onder andere doordat we deze gegevens veel gemakkelijker onthouden. In de sociale psychologie is dit bekend geworden als de beschikbaarheidsheuristiek (*availability* heuristiek); van dingen waarover we snel voorbeelden uit ons geheugen kunnen opdiepen, denken we ook dat ze vaker voorkomen. Een experiment dat dit aantoonde, is de vraag of er in het Nederlands meer woorden met een R beginnen dan met een R als tweede of derde letter.

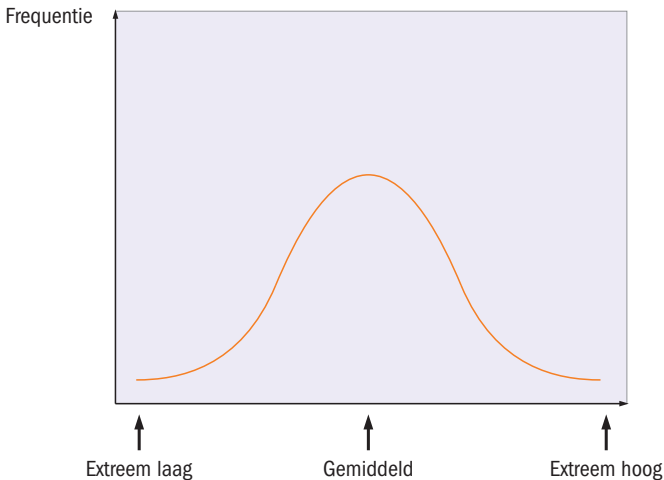
Beschikbaarheidsheuristiek

Omdat we sneller de eerste letter van een woord kunnen ophalen uit ons geheugen, denken we ook dat woorden die beginnen met een R vaker voorkomen, terwijl dit feitelijk niet het geval is.

1.3.4 Regressie naar het gemiddelde

'Regressie naar het gemiddelde' is een verschijnsel dat zelfs de meest doorgewinterde wetenschappers nog wel eens in de maling neemt. Wat regressie naar het gemiddelde is, kan het best verduidelijkt worden aan de hand van een aantal voorbeelden. Neem je eigen gezondheid. Soms voel je je fit, soms wat vermoeid en verkouden en enkele keren per jaar zul je je echt beroerd voelen. In het laatste geval heb je bijvoorbeeld een griep onder de leden, je bent vermoeid van het werk of je studie, je hebt een tijd niet goed geslapen en bent daarbij ook nog eens snipverkouden. Op het dieptepunt van je ellende komt een vriend langs en die zegt dat hij kan toveren. Hij spreekt een toverspreuk uit en beweert dat je je over twee weken een stuk beter zult voelen. En warempel, twee weken later voel je je inderdaad een stuk beter. Kan die vriend dus toveren? Nee, natuurlijk niet. Hij heeft je gefopt met regressie naar het gemiddelde.

FIGUUR 1.5 De normaalverdeling



Kijk eens naar figuur 1.5. Jouw gezondheid volgt een vergelijkbaar patroon. Doorgaans zul je je gemiddeld fit voelen. Die toestand komt gelukkig ook het vaakst voor. Een enkele keer voel je je superfit en ben je bijvoorbeeld in staat om in jouw sport je persoonlijke record te verbeteren. Een andere keer voel je je allerbelabberdst en zit je dus helemaal links in de curve. Gelukkig komt dit laatste niet zo vaak voor. Dit betekent dat wanneer je of op je best of op je slechtst bent je daarna hoogstwaarschijnlijk weer meer naar het gemiddelde zult trekken. Die kans is in ieder geval aanzienlijk groter dan dat je nog extremer zult scoren, dat je je met andere woorden nog zieker of nog fitter zult gaan voelen.

Een onderzoeker die een medicijn ontwikkelt, dit aan zieke mensen geeft en constateert dat het overgrote deel van die mensen er na zijn behandeling be-

ter aan toe is, en daaruit de conclusie trekt dat zijn nieuwe medicijn dus werkt, begaat een fout die uit regressie naar het gemiddelde voortkomt. Hij vergeet het effect van spontaan herstel. Hoe zou hij de werkzaamheid van zijn medicijn wel kunnen aantonen? Simpelweg door een controlegroep van even zieke mensen te nemen en die of geen medicijn of een nepmedicijn (een placebo) toe te dienen. Hij zal dan merken dat ook in deze groep een flink aantal mensen is opgeknapt. Hij mag alleen claimen een werkzaam nieuw medicijn te hebben ontwikkeld wanneer de groep die zijn medicijn heeft geslikt aanzienlijk sterker is vooruitgegaan dan de controlegroep die het placebo slikte.

Effect van
spontaan herstel

1

Dat extreme of bijzondere gebeurtenissen ons meer opvallen dan gemiddelde gebeurtenissen kan de diagnost parten spelen bij het interpreteren van bijvoorbeeld testgegevens. Extreme scores vallen meer op en wegen dus zwaarder mee. Beoordelaars kunnen dan in de valkuil stappen dat ze geen rekening houden met de regressie naar het gemiddelde.

Extreme scores

In psychologische testen is gelukkig wél rekening gehouden met dit fenomeen door een kenmerk, bijvoorbeeld extraversie, niet met één vraag of item te bevragen, maar er bijvoorbeeld vijf of tien vragen over te stellen, waarbij allerlei verschillende aspecten van de factor extraversie worden onderzocht. Mocht een cliënt dan één vraag of item zeer extreem beantwoorden, dan wordt dat weer rechtgetrokken door de andere items die op dezelfde factor laden. Maar dan nog kan het voorkomen dat een extreme totaalscore niet de werkelijke stand van zaken weergeeft. In paragraaf 2.2 zullen we dieper ingaan op de manier waarop testconstructeurs dit probleem hebben opgelost.

1.3.5 Eerste en laatste indruk

Een verschijnsel dat te maken heeft met de beschikbaarheidsheuristiek is het *primacy*- en *recency*-effect. Van een les, een toneelstuk, een film of een voetbalwedstrijd onthouden we het best wat er aan het begin en aan het eind gebeurde. In het begin zijn we nog fris en alert, terwijl het eind datgene is wat zich het meest recent heeft voorgedaan, waardoor we daarvan minder vergeten. Zaken die in het midden spelen, verdwijnen het snelst uit ons geheugen. We zagen al dat dingen die we makkelijk onthouden, vlot mentaal beschikbaar (*available*) zijn en het sterkst hun stempel drukken op onze meningsvorming. Dit gebeurt ook wanneer we bijvoorbeeld als personeelsfunctionaris een gesprek houden met een sollicitant of als hulpverlener een intake voeren met een potentiële cliënt. De eerste indruk weegt erg zwaar ('je krijgt nooit een tweede kans een eerste indruk te maken') én wat er aan het eind van het gesprek gebeurde blijft hangen. Hiermee moeten we als diagnosten rekening houden, bijvoorbeeld door gedurende het gehele gesprek aantekeningen te maken over wat de cliënt of sollicitant zegt en niet alleen op ons geheugen te vertrouwen, want dat is notoir feilbaar.

Primacy- en
recency-effect

Eerste indruk

1.3.6 Voorbarige reductie van cognitieve dissonantie

Een van de bekendste theorieën uit de sociale psychologie is de cognitieve dissonantiereductietheorie van Leon Festinger (1957). Hij ontdekte dat mensen op een zeer krachtige manier gedreven worden door een streven ideeën onderling en ideeën en hun acties met elkaar in overeenstemming te houden. We hebben moeite met botsende ideeën. Voor onszelf is het lastig te verkopen dat we enerzijds roken en anderzijds weten dat roken longkanker veroorzaakt, of dat we vinden dat de bio-industrie iets slechts is maar we wel kiloknallers vlees kopen. We zullen dan sterk gemotiveerd zijn deze dissonantie op te heffen, bijvoorbeeld door niet meer te roken of geen vlees meer

Cognitieve
dissonantie-
reductietheorie

te eten. Of we reduceren de dissonantie door onze overtuigingen over roken en vlees eten aan te passen. We vertellen onszelf bijvoorbeeld dat we wel roken, maar dat we verder supergezond leven, waardoor de kans op kanker toch heel klein is. Of dat je na het stoppen met roken na tien jaar weer een normale kans op longkanker hebt en je jezelf verteld hebt dat je stopt met roken wanneer je vijftig wordt. Ongeacht of je dit ook echt doet, is op die manier de cognitieve balans in je bewustzijn weer hersteld.

Experimenteel bewijs van cognitieve dissonantiereductie

Festinger onderzocht tal van situaties waarbij cognitieve dissonantiereductie een rol speelt en ontwikkelde experimenten om het fenomeen wetenschappelijk aan te tonen. Het bekendst daarvan is het 20 dollar-experiment uit 1957. Hierin deden twee groepen studenten een middag lang een vervelende klus. De ene groep kreeg daarvoor de karige vergoeding van 1 dollar, de andere kreeg 20 dollar (voor studenten in die tijd een behoorlijk bedrag). Enige dagen later werd de studenten gevraagd hoe leuk ze de taak gevonden hadden. En wat bleek? In tegenstelling tot wat de meeste mensen verwachten vonden de studenten die 1 dollar kregen de klus veel leuker dan de mensen die 20 dollar betaald werd. Vreemd! Beloning veroorzaakt kennelijk geen intrinsieke motivatie. Hoe komt dat?

Festinger verklaart deze uitkomsten door te wijzen op de cognitieve dissonantie van de groep die 1 dollar kreeg. De studenten die in deze groep zaten, hadden kunnen stoppen met de klus, of ze hadden vooraf al kunnen weigeren om aan de klus te beginnen, of hadden er halverwege de brui aan kunnen geven of kunnen protesteren tegen de karige beloning. Dit hebben ze allemaal niet gedaan. Dus moesten ze op een andere manier goedpraten dat ze het werk toch hebben uitgevoerd. Sommigen deden dit door te onderstrepen dat ze de onderzoeksleider een aardig iemand vonden en hem niet wilden teleurstellen, maar de meesten maakten zichzelf wijs dat de klus eigenlijk helemaal niet zo vervelend was. Of sterker nog, best leuk was eigenlijk. De 20 dollar-groep heeft geen last van deze cognitieve dissonantie. Voor hen was het gewoon een vervelende klus die ze alleen maar uitvoerden omdat ze er goed voor werden betaald.

Vasthouden aan bestaande opvattingen

De neiging om onze eerste indruk erg zwaar te laten wegen komt voort uit processen die met cognitieve dissonantiereductie te maken hebben. Latere informatie gaan we selectief filteren, of dat wat met ons eerste idee overeenkomt accepteren we te gemakkelijk en dat wat ermee in strijd is poetsen we weg en zien we liever niet. We zoeken naar een logisch samenhangend beeld. De zogenoemde horn- en halo-effecten zijn op dit laatste gebaseerd. Een halo is een stralenkrans. In de psychodiagnostiek spreken we van een halo-effect als we op basis van een aantal positieve zaken die we van iemand hebben gezien (bijvoorbeeld dat hij hard traint en bijzondere prestaties in de sport levert) aannemen dat hij ook op andere vlakken (bijvoorbeeld in zijn studie of werk) bijzonder zal presteren. En omgekeerd, wanneer we horen dat iemand aan drugs verslaafd is geweest, nemen we bijna automatisch aan dat hij als student of werknemer ook wel onbetrouwbaar en lui zal zijn, terwijl dit verband helemaal niet aanwezig hoeft te zijn.

Professor Bavelas deed eens een zeer belangwekkend experiment waarbij proefpersonen aan de hand van foto's het onderscheid moesten leren tussen gezonde en zieke lichaamscellen (Watzlawick, Beavin & Jackson, 2009). Bij de ene groep gaf Bavelas realistische feedback. De leden van

deze groep leerden vrij snel het onderscheid te maken. Bij een tweede groep proefpersonen gaf hij verkeerde feedback. Deze proefpersonen konden dus onmogelijk het verschil ontdekken tussen zieke en gezonde cellen. Toch meenden deze proefpersonen zelf het na verloop van tijd wél te kunnen. Zij construeerden daartoe zeer ingewikkelde (maar helaas onjuiste) theorieën over het verschil tussen zieke en gezonde cellen. Bovendien bleken ze veel fanatieker in het verdedigen van hun theorie dan de leden van de eerste groep. Kennelijk hebben de leden van de tweede groep er zo hard voor moeten werken, dat ze hun theorie niet zomaar prijsgeven.

Ook als psychodiagnost zijn we voortdurend theorieën aan het ontwerpen over cliënten. Wanneer je bijvoorbeeld een aantal symptomen hebt waargenomen en steeds meer in het idee bent gaan geloven dat een cliënt last heeft van faalangst, dan loop je het risico dat je informatie die dit beeld lijkt te bevestigen te zwaar laat wegen en informatie die het kan ontkrachten het liefst onder het tapijt veegt. In hoofdstuk 5 zullen we zien welke procedurele stappen er zijn ontworpen om deze riskante maar o zo menselijke fout te voorkomen.

Gehechtheid aan eigen oordelen

1

1.4 Remedies tegen beoordelingsfouten

Je kunt de systematische psychodiagnostiek beschouwen als één grote poging om menselijke beoordelingsfouten tegen te gaan. Testen en andere formele methoden zijn de structurele kant hiervan; hoe dat precies werkt komt later in dit boek aan de orde. Je kunt oplossingen in testen inbouwen, maar daarnaast is het belangrijk om het menselijke beoordelingsproces in rechte banen te leiden.

Eén daarvan is de eis om te zoeken naar *Contrary Evidence*. Dit betekent dat als jij je mening hebt gevormd, bijvoorbeeld: Jantje is niet geschikt om naar het vwo te gaan, je dan bewust gaat zoeken naar aanwijzingen voor de stelling dat Jantje daar juist wel geschikt voor zou zijn. Lukt je dit, dan moet je je mening bijstellen. Heb je het echt geprobeerd en lukt het desondanks niet, dan kun je met iets grotere zekerheid zeggen dat je aanvankelijke mening klopte. In de wetenschapsfilosofie heet dit falsificatie. Dit komt erop neer dat je experimenten bedenkt die een wetenschappelijk idee onderuit kunnen halen (in plaats van krampachtig nog meer ondersteunend bewijs voor je theorie bij elkaar te schrapen). Als het niet lukt een theorie onderuit te halen, dan is dit een bewijs voor de kracht van die theorie (Popper, 1972).

Contrary Evidence

Falsificatie

Een andere methode om eenzijdigheid in je meningsvorming tegen te gaan, is meerdere personen in te schakelen om tot een eindoordeel te komen, de zogenoemde multi-ratermethode. Bijvoorbeeld: de toegepast psycholoog interviewt een kind, zijn collega heeft een gesprek met de ouders, een test-assistent neemt een aantal testen af en de schooljuffrouw observeert het kind aan de hand van een gestructureerd observatieschema (een lijst waarop zij kan aankruisen hoe vaak en hoe intensief bepaald gedrag plaatsvindt, bijvoorbeeld samen spelen, ruzie maken, anderen helpen et cetera). Al deze gegevens worden verwerkt volgens een vooraf vastgesteld protocol. Dit protocol is gebaseerd op wetenschappelijk onderzoek. Bij dergelijke gestandaardiseerde, statistische beslissingsprocedures wordt dus een deel van het besluitvormingsproces van mensen afgenomen en overgezet in een computer of een precies omschreven procedure.

Multi-ratermethode

Samenvatting

- ▶ Psychodiagnostiek maakt integraal onderdeel uit van onze moderne maatschappij. Voortdurend worden beslissingen genomen over mensen, door henzelf of door anderen. De professionele psychodiagnostiek kan bijdragen aan de kwaliteit van deze beslissingen.
- ▶ Beslissen is altijd voorspellen. De kwaliteit van een voorspelling kan bepaald worden door de voorspelling met de uiteindelijke gebeurtenissen te vergelijken. Op basis hiervan kan een beslissing getypeerd worden als een Valid Positive, een Valid Negative, een False Positive of een False Negative.
- ▶ Wanneer de weerman regen voorspelt en het gaat ook regenen, dan is dit een Valid Positive. Gaat het niet regenen, dan is dit een False Positive. Voorspelt hij droog weer en blijft het ook droog, dan is dit een Valid Negative. Heeft hij droog weer voorspelt, maar gaat het toch regenen, dan is dit een False Negative.
- ▶ Het begrip correlatiecoëfficiënt geeft ons een instrument in handen om het verband tussen twee variabelen uit te drukken in een exact getal. De correlatiecoëfficiënt is een zeer veel gebruikt begrip in de psychodiagnostiek. We komen het onder andere tegen wanneer we de betrouwbaarheid en validiteit van een test in een cijfer willen uitdrukken.
- ▶ Het menselijke beoordelingsvermogen is niet onfeilbaar. Uit onderzoek zijn verschillende veelvoorkomende valkuilen naar voren gekomen, zoals het overschatten van specifieke kansen, de beschikbaarheidsheuristiek, het negeren van regressie naar het gemiddelde, het te veel gewicht toekennen aan de eerste en laatste indruk, en de voorbarige reductie van cognitieve dissonantie.
- ▶ Psychodiagnostiek kan worden opgevat als een systematische aanpak van beoordelingsfouten bij het inschatten van eigenschappen van mensen. Deze preventieve maatregelen komen we tegen in de instrumenten en procedures die de psychodiagnostiek voorschrijft, maar is ook voor een belangrijk deel gebaseerd op de manier waarop de diagnost werkt; in zijn attitudes, vaardigheden en zijn voortdurende focus beoordelingsfouten te voorkomen.