

Samenvatting psychodiagnostiek 1

Toegepaste psychologie

1-1-2012

Lessius, Antwerpen

Ulrike

Hoofdstuk 0: Het psychodiagnostische proces

Klinische psychodiagnostiek is een professionele activiteit die steunt op empirische kennis. Psychologische diagnostiek verwijst naar het verzamelen van informatie en het nemen van beslissingen op basis van deze informatie door een geschoolde diagnosticus. Het psychodiagnostische proces verwijst naar de manier waarop deze informatie verzameld moet worden. De psychodiagnosticus moet bij het verzamelen van informatie rekening houden met 3 standaarden:

1. Ethische standaarden
2. Sociale standaarden
3. Methodologische standaarden

Psychodiagnostiek betekent...

Beslissen

< Grieks (onderscheiding): onderscheidings-, beslissingsvermogen

“een systematisch beslissingsproces” (De Bruyn, 1988)

Vaststellen van individuele verschillen

“een onderzoekstaak om individuele verschillen tussen mensen vast te stellen en om deze kennis toe te passen op het enkele individu” (De Zeeuw, 1983)

Het vaststellen van individuele verschillen gebeurt adhv meetinstrumenten

Verzamelen van info over persoon en zijn/haar omgeving

“op wetenschappelijk verantwoorde wijze verzamelen van info omtrent persoon en zijn omgeving met het oog op het nemen van beslissingen” (Van Aarle, 1990).

Hulp bieden

“psychodiagnostisch proces is het proces van hulp bieden aan personen met hun vragen en problemen” (Walsh & Betz, 2001)

Diagnostiek mondt uit in behandeling/interventie Zie vb1

Beschrijven van ontwikkelingsniveaus (zwakke én sterke kanten) van cliënt

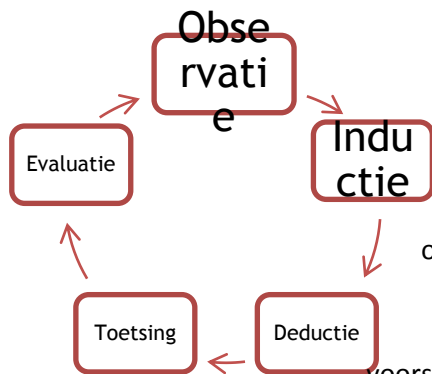
“developmental assessment op diverse ontwikkelingsgebieden zoals cognitie, motoriek, socio-emotionele vaardigheden, ...” (Johnson & Sheeber, 1999) Zie vb1 en vb2

Testen en diagnostisch proces

Psychometrie houdt zich bezig met de formeel-statistische fundamenteën van het meten en valideren van individuele verschillen. De toegepaste psychometrie focust zich op het ontwikkelen van de tests en het onderzoeken van de betrouwbaarheid en validiteit van deze testen.

Het proces

Het psychodiagnostisch proces wordt gereguleerd door de empirische cyclus en bevat volgende aspecten:



1. Observatie: is het verzamelen van informatie en groeperen van empirisch materiaal, waaruit gedachten over de totstandkoming en het voortduren van probleemgedrag gevormd worden.

2. Inductie: omvat de formulering van de theorie en hypothesen over het gedrag.

3. Deductie: is het afleiden van toetsbare voorspellingen.

4. Toetsing: aan de hand van nieuw materiaal wordt er nagegaan of de voorspellingen juist zijn.

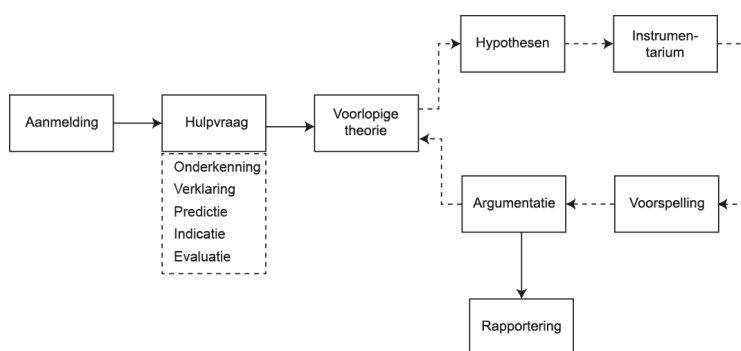
5. Evaluatie.

Oordeelfouten

Onderzoek naar klinische oordeelfouten komen in 3 psychologie domeinen voor:

1. Het psychometrie: Een klinische predictie is altijd minder accuraat dan een predictie gebaseerd op formele beslissingsregels.
2. De cognitieve psychologie: Cognitieve heuristieken die mensen in hun leven gebruiken werken vrij goed, maar deze heuristieken worden niet correct aangewend wanneer mensen moeten omgaan met probabilistische informatie.
3. De sociale psychologie: de clinicus die in contact is met de cliënt is geneigd om informatie te selecteren en interpreteren die overeenkomen met zijn/haar causale attributies, stereotypen en karakteristieken.

modelleren van het proces

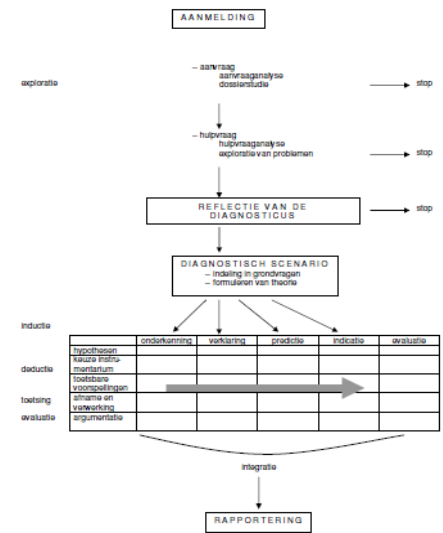


- Statistisch predictiemodel: De gestandaardiseerde scores op verschillende testen worden overeenkomstig in de mate waarin ze een invloed hebben op een bepaald criterium op zo een manier gewogen dat de predictie van het criterium

optimaal is.

Samenvatting psychodiagnostiek 1

- Utiliteitsmodel: Het beslissingsproces wordt geformaliseerd als combinatie van de probabiliteit en de waarden van verschillende keuzealternatieven.
- Algoritmisch model: lost de nadelen van de vorige 2 modellen op, maar is zelf ook niet 100% betrouwbaar.



Het diagnostisch model

In de eerste contactgesprekken analyseert de diagnosticus zowel de hulpvraag van de cliënt als de aanvraag van de verwijzer of opdrachtgever. De meeste vragen van de cliënten, verwijzers, opdrachtgevers en psychodiagnostici kunnen herleid worden tot 5 basisvragen:

1. **Onderkenning:** Wat zijn de problemen, wat lukt er nog en wat gaat er mis?
 - a. Inventarisatie en beschrijving van het probleemgedrag.
 - b. Ordening en categorisering van probleemgedragingen in disfunctionele gedragsclusters of stoornissen.
 - c. Inschatting van de ernst van het probleemgedrag.
2. **Verklaring:** Waarom zijn bepaalde problemen er en waarom blijven ze bestaan?
Een verklaring geeft antwoord op de vraag waarom er een probleem is en bevat volgende aspecten:
 - a. Het deelprobleem.
 - b. De condities die het optreden van een probleem verklaren.
 - c. De relatie tussen a en b.
3. **Predictie:** Hoe gaan de problemen van de cliënt zich in de toekomst verder ontwikkelen?
4. **Indicatie:** Hoe kunnen de problemen verholpen worden?
Heeft de cliënt behandeling nodig en zo ja, welke behandeling is er dan het meest geschikt?
Hierbij komen 3 elementen bij kijken:
 - a. Kennis over behandelingen en behandelaars.
 - b. Kennis over het relatieve nut van behandelingen.
 - c. Kennis over de aanvaarding van de indicatie door de cliënt.
5. **Evaluatie:** Zijn de problemen door de interventie afdoende verholpen?
 - a. Is er in de therapie rekening gehouden met de diagnose en behandelingsmodel?
 - b. Heeft het proces en de behandeling de verandering van gedrag bewerkstelligd?

Het toetsen van een voorlopige theorie vereist 5 diagnostische handelingen:

1. De voorlopige theorie wordt omgezet in concrete hypothesen.
2. Een specifiek onderzoeksinstrument wordt gekozen dat een antwoord kan geven op de hypothesen die gesteld worden.
3. Er worden voorspellingen gedaan over resultaten of uitkomsten op dit instrumentarium zodat men duidelijk op voorhand weet wanneer men de hypothese zal aanvaarden of verwerpen.
4. Men neemt de instrumenten af en men verwerkt de resultaten.

5. Op grond van de verkregen resultaten aanvaardt of verwierpt men op een beargumenteerde wijze de hypothesen.

Rol van psychodiagnosticus

Omschrijving

BaTP kan zowel zelfstandig als in teamverband op kritische wijze een psychodiagnostisch stappenplan ontwikkelen en uitvoeren waaronder exploratie van de aanvraag, probleemanalyse en het genereren van hypothesen, het opstellen en uitvoeren van de onderzoeksstrategie, integratieve verwerking van de resultaten, rapporteren met aanknopingspunten voor indicatiestelling. Deze rol is een van de twee belangrijkste kernrollen, de andere is de rol dienstverlener.

Competenties

- Informatie verzamelen
- Analyseren van de vraagstelling
- Opstellen en toepassen van een onderzoeksstrategie
- Integratief werken
- Rapporteren

Informatie verzamelen

BaTP verzamelt op zelfstandige wijze relevante info via wetenschappelijk onderbouwde methoden en brengt vanuit een gepast psychologisch referentiekader de ingewonnen info met elkaar in verband.

Analyseren van de vraagstelling

BaTP analyseert een psychosociale vraag en/of probleem in zijn componenten op een methodische wijze. De BaTP herformuleert de vraagstelling in toetsbare hypothesen.

Opstellen en toepassen van een onderzoeksstrategie

BaTP kiest, in functie van de vraagstelling, de meest geschikte psychometrisch onderbouwde en/of niet psychometrische methode(n) en past die toe om de vooropgestelde hypothesen te toetsen.

Integratief werken

BaTP ordent en interpreteert de verkregen diagnostische gegevens om zo tot een objectieve synthese te komen en er een psychologisch besluit en/of advies uit af te leiden.

Rapporteren

BaTP verwoordt het besluit en/of advies en het proces dat daartoe heeft geleid op een heldere manier naar collega's en cliënten.

Hoofdstuk 1: Historische ontwikkeling van het testen

Enkele historische voorlopers

Al heel vroeg in de tijd zijn er voorbeelden gevonden van vormen van testen. Zelfs bij meer dan 1000jaar terug zie je prille vormen van testen terug. Reeds voor 2000 v. CHR liet een Chinese keizer zijn dienaren eens per 3 jaar onderzoeken met een soort van vorderingstoets, op basis waarvan zij werden gepromoveerd of ontslagen. Ook in het Oude Testament is er een voorbeeld te vinden van selectie met behulp van een persoonlijkheidstest. Gideon reduceert zijn 32 000 man sterke leger eerst op basis van een soort zelfbeoordeling, waarbij hij de 'angstigen' laat vertrekken. In de Middeleeuwen en de nieuwere tijd komt men herhaaldelijke pogingen tegen om intellectuele of karaktereigenschappen te beschrijven en te meten met behulp van zuiver fysieke kenmerken van gelaat, schedel of lichaam en zelfs buitenaardse determinanten. De meeste school- en academische examens geschieden vroeger mondeling, hoewel ze aan sommige universiteiten zoals Bologna en Leuven, reeds in de middeleeuwen een sterk formeel karakter hadden.

Testen zijn ontstaan van uit een maatschappelijk, praktisch probleem. Er waren belangrijke invloeden (Gebeurtenissen, personen, ...) die een rol speelden bij de testontwikkeling en psychologische diagnostiek.

De meer officiële testdiagnostiek is ontstaan in de 19^{de} eeuw, maar je vind al heel ver terug voorlopers van de testbeweging die leidde tot procedures voor bepalen van individuele verschillen in intelligentie en persoonlijkheid. Maar de geschiedenis van de diagnostiek is de geschiedenis van de:

- Testen.
- Testtheorie.
- Psychologische theorievorming.
- Diagnostisch proces.
- ...

Geschiedenis diagnostiek vooral over opkomst, verspreiding en gebruik van psychologische tests en meetinstrumenten behoort tot de geschiedenis van testbeweging.

Periode tot Binet-Simon test

Bijdrage van de psychiatrie



Er waren inzichten van de psychiatrie in de aard van krankzinnigheid en zwakzinnigheid. Pinel (1793) is een Parijse arts en heeft als eerste een interesse getoond in de psychiatrische stoornissen. Hij was ook vernieuwend omdat hij de eerste was die een andere aanpak wou, een menselijke manier. Hij is directeur geweest van het ziekenhuis Bicêtre. Hier werden mannen geplaatst, omdat ze niet gewenst waren. Deze mannen konden niet meer voor zichzelf zorgen, waren gehandicapt, deden aan criminaliteit en konden hun emotionele stoornissen niet meer aan. In het 2^{de} ziekenhuis La Salpêtrière zaten alleen maar vrouwen. Ook deze vrouwen waren verschillende soorten patiënten. Maar ook

kinderen.

Pinel heeft een aantal goede dingen gedaan, zijn aanpak was vernieuwend. Hij ging de mensen op een menselijke manier benaderen en dat er verschillende activiteiten werden gepland, gepaste voeding, goede fysiek, enz. Ten tweede vond hij het belangrijk dat zijn personeel geschoold werd. Hij vond het belangrijk om de patiënten te scheiden en een eigen document van elke patiënt aan te maken.

Een student van Pinel, Esquirol (1838), heeft ook in Salpêtrière gewerkt. Esquirol heeft een onderscheid gemaakt tussen zwakzinnigheid en psychiatrische stoornis. Hij zei dat zwakzinnigheid permanent en onomkeerbaar is en ongeneeslijk is. Terwijl psychiatrische stoornis op latere leeftijd kon ontstaan en dat deze stoornissen kunnen verbeteren. Ten tweede heeft hij ook al een vorm van classificatie gezien binnen zwakzinnigheid (idiotie). Hij zag hier 3 types in op basis van de taal.



Séguin (1837) is een student geweest van Esquirol, maar hij ging nog een stap verder. Hij was van mening dat idiotie een degeneratie was van het centraal zenuwstelsel, en dat kon wel gestimuleerd en getraind worden. Hij was begonnen met de eerste behandelingen/therapieën. De training had voornamelijk betrekking op de motorische en sensorische functies. Het is dan ook niet verwonderlijk dat de 'test' die Séguin ontwikkelde op deze motorische en sensorische functies betrekking had. Het ging hier om een type test dat later 'performance test' zou gaan heten. Deze test is bekend als het Séguin Form Board, waarvan later diverse aanpassingen werden gepubliceerd. In psychiatrische kringen in Frankrijk bleef aandacht voor de geestelijke onvolwaardigheid bestaan. Bekende namen zijn Charcot, Janet en Ribot, en uiteindelijk Alfred Binet, die reeds in zijn jonge jaren geboeid was door het verschijnsel zwakzinnigheid.

Ook in Duitsland hielden psychiaters zich bezig met het onderzoek van diverse geestelijke vermogens. In 1885 werd reeds door Rieger een voorstel gedaan voor een algemeen toepasbare methode voor het intelligentie onderzoek.



Ook Kraepelin (1883) was een Duitse arts die interesse had in psychiatrische stoornissen. Hij heeft een compendium der psychiatrie proberen te beschrijven. Dat was de eerste classificatie in de psychiatrie. Hij zag in dat een psychiatrische stoornis niet uit een symptoom bestaat, maar uit meerdere symptomen. Hij heeft die symptomen dan gevormd tot symptoomclusters. Maar hij zag ook in dat niet een symptoom niet bij een bepaalde stoornis komt, maar bij meerdere stoornissen. Kraepelin zag een fysiologische oorzaak voor die psychiatrische stoornissen. Dit wilde zeggen dat het ook behandelbaar was. Ook meting bleek noodzakelijk te zijn. Om stoornissen te kunnen classificeren, moet je ze ook kunnen meten.

In 1897 probeerde Ebbinghaus psychologische vermoeidheid te meten met een test waarin onvolledige zinnen moesten worden aangevuld, maar die in werkelijkheid functioneerde als een intelligentietest. In datzelfde jaar deed Ziehen het voorstel om de evaluatie van de antwoorden niet reeds vooraf te verrichten, maar deze te laten afhangen van de prestaties van de geteste personen.



Bijdrage experimentele psychologie

De experimentele psychologie werd vooral in Duitsland beoefend, en bestaat ook al van de 19^{de} eeuw. De experimentele psychologie is geïnteresseerd in de algemene wetmatigheden, in verschillen tussen conditie. Opvallend is echter dat de experimentele psychologie naast een stimulerende ook een remmende werking had op de testpsychologie. De stimulans lag in de waarde die gehecht werd aan de exacte beschrijving van experimenteercondities, de rigoureuze controle van de variabelen en de nauwkeurige verwerking van de uitkomsten. Ironisch genoeg, tevens de remmende invloed van de experimentele psychologie op testpsychologie. De experimentator vond de verschillen tussen proefpersonen eerder vervelend dan boeiend. Hij schreef ze toe aan fouten in plaats van aan werkelijke verschillen tussen proefpersonen. Dit laatste is nu datgene waar het bij testen om gaat.

Wundt (1879) richtte het eerste laboratorium op. Het was een start van systematisch experimenteel onderzoek op grote schaal. Wundt was geïnteresseerd in de denksnelheid door problemen die in zijn tijd plaatsvonden. Ebbinghaus was degene die zich interesseerde voor het geheugen. Hij deed proeven om te zien wat mensen konden onthouden en hoelang, de verschillende resultaten bij verschillende mensen. Met enorme grondigheid, volledigheid en ook spitsvondigheid werden talloze functies van de mens onderzocht, maar ook in dit werk bleven twee obstakels voor de ontwikkeling van het testen bestaan. Ten eerste was het doel van het onderzoek de generaliseerbaarheid van de wetten en samenhangen, terwijl afwijkingen en verschillen werden beschouwd als experimenteerfouten. Ten tweede bleef het onderzoek beperkt tot de primaire sensorische en motorische functies, en werden de hogere en meer complexe cognitieve en intellectuele processen goeddeels buiten beschouwing gelaten.



Toch ging men op een zeker moment anders aankijken tegen individuele verschillen. Tegen de zin van Wundt, promoveerde bij hem een Amerikaanse experimentator, Cattell, op individuele verschillen in reactietijd. In 1890 publiceerde Cattell een artikel in het tijdschrift Mind waarin voor de eerste maal het woord 'test' werd gebruikt, en waarin hij een reeks van deze mentale (psychofysische) tests besprak en voor verder onderzoek aanbeval. Terug in de Verenigde staten stichtte Cattell een laboratorium op voor experimentele psychologie en testpsychologie.

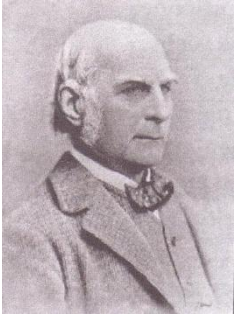
Een tweede iets was de ontwikkeling van meetinstrumenten, dit werd door de experimentele psychologen verder ontwikkeld:

1. Tachistoscoop (1903): Toestel waarmee visuele stimuli kunnen worden aangeboden voor een zeer korte tijdsspanne (msec)
2. Chronograaf.
3. Bewegingsmodellen van het oog.

Een derde is de manipulatie van onderzoekcondities. De storende factoren moest je zoveel mogelijk proberen te elimineren of constant houden om zo nauwkeurig mogelijk te schatten. Het gestandaardiseerd werken komt vanuit de experimentele psychologie in de tijd van Wundt. Tenslotte leerde we van de experimentele psychologen dat we hypothesen gebruiken en toetsen, en zo ontstond het diagnostisch proces.

Het had ook nog beperkingen. Er was de focus op algemene wetten en samenvattingen en niet op afwijkingen en verschillen. Ook de focus op primaire sensorische en motorische functies en niet op hogere en complexere cognitieve en intellectuele processen.

Bijdrage van de genetica



Galton was een Brit die een zeer brede interesse had. Zijn overtuiging was van het Engels empirisme. Kennen was enkel mogelijk via de zintuigen. Daarnaast was hij ook nog obsessief bezig met het meten. Galton was geïnspireerd door Darwin. Galton wilde aantonen dat psychische eigenschappen ook erfelijk zijn en dezelfde wetten volgen als lichamelijke eigenschappen. Galton gebruikte eugenetica naar het fokken van dieren die het meest succesvol leken. Metingen zouden als basis kunnen dienen om de besten te selecteren.

Voor het bewijzen dat psychische eigenschappen ook erfelijk waren, kwamen er drie elementen voor in zijn onderzoek:

1. Wenselijkheid van het onderzoek van individuele verschillen.
2. Systematisering van de onderzoekstechnieken.
3. Resultaten van de onderzoeken van Galton uit te drukken in termen van afwijkingen van het gemiddelde.

Galton deed die selecties via psychofysische metingen voor het mentale vermogen. Metingen van het mentale vermogen gebeurde via psychofysische metingen van elementaire functies. Hij deed dit omdat elementaire functies eenvoudig en exact te meten zijn en ze zouden weinig beïnvloed worden door milieu en opvoeding, zijn representatief voor het mentale functioneren. Hij was de ontwikkeling van het gebruik van objectieve testen en maakte gebruik van gestandaardiseerde procedures. Galton maakte ook gebruik van wiskundige technieken. Hij maakte gebruik van co-relations en regressie - analyse.

Een jongere collega van Galton, Pearson, werd door deze verwerkingsmogelijkheden geïnspireerd en heeft vervolgens zeer veel betekend voor de statistiek en indirect voor de testpsychologie. Aan Pearson danken wij de correlatiecoëfficiënt, de multipele correlatie, de factoranalyse en verscheidene andere nuttige statistische technieken.

De belangstelling van Cattell (1860-1944 in V.S.) voor individuele verschillen kan ook teruggevoerd worden op de genetica. Hij heeft een invloed gehad op de verdere ontwikkeling van testen. Er was een link tussen de Engelse traditie en Amerikaanse behoeften aan tests. Mentale tests weerspiegelen de werkwijze van Galton: handgrip, discrimineren gewichten, enz. de overtuiging was dat mentale en lichamelijke energie niet van elkaar konden onderscheiden worden.



Cattell heeft een testbatterij ontwikkeld. Dit werd op een jaarlijkse schaal afgenomen bij schoolkinderen en studenten met oog op het bepalen van intelligentie en het voorspellen van schoolprestaties. Wissler, een student van Cattell, heeft zich toegelegd op de testbatterij. Hij onderzocht of de scores op de testbatterij schoolresultaten van studenten kunnen voorspellen en in hoeverre de verschillende proefjes onderling samenhangen. Kortgezegd of dat de testbatterij valide was.

Periode tussen Binet-Simon-test en WO-I

Binet (1857-1911) is een veelzijdige en zeer actieve persoon. Hij wees de sensorische tests voor onderzoek naar intelligentie af. Er zijn individuele verschillen die je moet bepalen aan de hand van hogere mentale processen. Hij stelde testen voor zoals geheugentests, verbeelding met behulp van inktvlekken, aandachtsonderzoek, enz. In 1905 krijgt Binet de kans om een test te ontwikkelen. Dit omdat er een verzoek van Ministerie van Onderwijs die de schoolplicht invoerde. Maar veel jongeren haakte af en ze wilden weten of dit kwam door de intelligentie of andere factoren. Daarom ontwikkelde Binet een test voor selectieve maatregelen voor kinderen die wel/niet het gewone onderwijs konden volgen aan de hand van 30 proefjes.



De kenmerken van deze test was:

- Niet bedoeld om alle aspecten van intelligentie te meten maar wel als maat voor algemene mentale ontwikkeling van de kinderen → totaalscore.
- Het was een zeer korte en praktische test.
- Het werd aantrekkelijk gemaakt zodat de kinderen hun aandacht er bij konden houden.
- Intelligentie wordt beschouwd als oordeelsvermogen en niet als lagere vaardigheden → accent op complexe mentale processen.
- Er waren nog geen gestandaardiseerde instructie.

Er volgen nog een aantal herwerking van de eerste versie. De Binet-Simon test werd afgenomen bij verschillende kinderen en werd per leeftijdsgroep nagegaan hoeveel een item kinderen goed beantwoorden. Dit was de mentaal niveau of de mentale leeftijd. De bepaling van het mentale leeftijd was de bepaling van het basale mentale niveau ; en voor elke 5 gelukte proeven boven het basale niveau wordt een extra jaar toegerekend.

Claparède (1924) bouwde voort op de ideeën van Binet en een groot deel van de activiteiten in zijn in 1912 gestichte Jean-Jacques Rousseau Instituut was gewijd aan de ontwikkeling en toepassing van tests. Bobertag vertaalde de Binet-Simon test in het Duits.



Stern (1911) ging een stap verder en gaat de mentale leeftijd vergelijken met de chronologische leeftijd. Zo is hij het IQ gaan bepalen. Terman bracht verandering in het concept van IQ, hij ging het IQ vermenigvuldigen met 100 om breuken te vermijden. Naast de nieuwe scoringswijze waren er veel vertalingen, Terman maakte een aantal Amerikaanse versie. Dit heeft geleid tot een nieuwe scoringswijze, maar heeft ook standaardinstructies en normen ontwikkeld. Dit werd dan de Stanford-Binet test.

Terman besteedde een groot deel van zijn werkzame leven aan de Amerikaanse versie van deze test, sindsdien algemeen bekend onder de naam Stanford-Binet. Reeds bij de eerste versie van 1916 was aan twee zeer belangrijke eisen voldaan: er werden standaardinstructies en normen geformuleerd. Ook nam Terman het idee van Stern over om de mentale leeftijd te vermenigvuldigen met 100.

Op basis van zijn theoretische onderzoeken kwam de Engelse onderzoeker Spearman tot de conclusie dat in alle tests twee intelligentiefactoren een rol speelden. Een algemene (g-)factor en een specifieke (s-)factor. Hiermee had hij zijn bekende twee-factorentheorie, die later door de Amerikaan Thurstone met diens multi-factorentheorie zou worden bestreden.



Intelligentie was volgens Spearman de gemeenschappelijke factor die men kan extraheren uit een willekeurige serie van intelligentietests. Doordat Binet steeds had gestreefd naar een breed scala van opgaven had hij daarmee onbewust een test voor de g-factor geconstrueerd.

Begin WO-I tot WO-II

Reeds voor WO-I had de Amerikaanse psycholoog Knox geëxperimenteerd met opgaven die niet verbaal van aard waren en een zogenaamd performance-karakter hadden. In 1917 publiceerden Pinter en Paterson een volledig niet verbale intelligentietest.



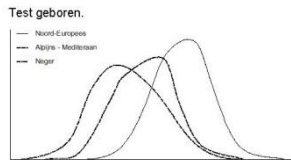
Goddard (1866-1957) deed een poging om de testen in het Engels te vertalen, maar deed dit met bepaalde bedoelingen. Hij was zeer sterk geïnteresseerd in zwakbegaafdheid. Hij was er van overtuigd in het begin dat je met de nodige training de zwakzinnigheid kon wegwerken. Maar later zag hij in dat er ook nog een erfelijk deel in meespeelt. Hij gebruikte beledigende termen om mensen te classificeren. En hij heeft ook onderzoek gedaan naar de prevalentie van mentale achterstand.

Goddard is ook bekend van de grootschalige testen. Hij organiseerde testings van immigranten op Ellis Island in de V.S. De problemen was dat de oorspronkelijke Franse test dat vertaald was in het Engels en daarna nog in andere talen. De immigranten waren zwaar vermoeid bij de afname van de test en de scores werden geïnterpreteerd met Franse normen. Door deze problemen is Goddard tot de conclusie gekomen dat de meeste immigranten zwakzinnig waren. Er was een positieve ontwikkeling. Ze beseften dat het vertalen op eigen houtje niet goed was. Ze ontwikkelde dan non-verbale tests voor immigranten.

Door WO I ontstond er een noodzaak om grote groepen personen voor uiteenlopende functies te selecteren. Dit werd in Amerika gedaan door de Army alpha test, een schriftelijke collectieve test; en een Army Beta test, een niet-verbale groepstest. Het probleem van de groepstests was dat er vele rekruten 0 scoorden op deze test. Dit kwam niet omwille van de mentale achterstand, bij deze test waren er ook wel fouten gebeurd.

Het gebruik van de test ging vaak vooraf aan de theorie, in plaats van dat men te werk ging volgens de gewenste omgekeerde volgorde. Wel stimuleerde Thurstone de aandacht voor kritische evaluatie van de test zelf. Hij was van mening dat de resultaten van een test betrouwbare testcores dienen op te leveren, en als men met een test gedrag buiten de testsituatie probeert te voorspellen dient de relatie tussen test en criterium van tevoren te zijn aangetoond. Hierdoor ging de statistiek een belangrijke rol spelen in de testpsychologie en werd de kwantificeerbaarheid van de testprestatie een noodzakelijke voorwaarde voor nuttig testgebruik.

Naast de formeel-statistische ontwikkeling werd door Kelley, Thurstone, Guilford, Thomson en Holzinger en Harman ook aan het onderzoek naar intelligentiestructuur een nieuwe dimensie toegevoegd. Zij introduceerden naast de reeds door Spearman gesignaleerde algemene en specifieke factoren, de zogenaamde groepsfactoren. De door Thurstone geïdentificeerde factoren: "Verbale comprehensie, woordvloeiing, getal faciliteit, ruimtelijke visualisatie, associatief geheugen, perceptuele snelheid en redeneren" nog steeds belangrijke onderscheidingen in de intelligentie-theorie en spelen een belangrijke rol in de testbatterijen voor intelligentie.



Brigham heeft op basis van deze resultaten om verschillende bevolkingsgroepen te vergelijken. Hij kwam tot de conclusie dat rasvermenging onherroepelijk zal leiden tot aftakeling van Amerikaanse intelligentie. Maar later herroept Brigham zijn uitspraken.

Naast de intelligentietesten zijn er ook nog andere soorten testen. Ook persoonlijkheidstesten zijn ontwikkeld achter maatschappelijke groei. Er was een nood aan instrumenten om emotioneel instabiele recruten te identificeren en dit gaf aan tot de 1^{ste} persoonlijkheidstest. Woodworth ontwikkelde hierdoor de Woodworth Personal Data Sheet. Dit was een schaal op basis van 116 symptoombeschrijvingen. Het nadeel van deze test was dat je de intentie van de test kon achterhalen waardoor een individu zich anders kon voordoen. Dit was een basis voor latere vragenlijsten. Een andere test is de MMPI. Hier was een koppeling van onaangepastheid aan psychiatrische categorieën. Bij deze test was er sprake van verschillende klinische- en validiteitsschalen. Bij deze test kon je de sociale wenselijkheid achterhalen wat niet kon bij de test van Woodworth.

Naast de persoonlijkheidsvragenlijsten heb je ook nog projectieve technieken. Een voorbeeld hiervan is de Rorschachtest. Rorschachs (1921) publiceert over ervaringen met inktvlekken en legt een verband tussen het duiden van inktvlekken en persoonlijkheid. Het probleem was niet bedoeld om bij de normale populatie af te nemen, maar bij zijn patiënten. Een andere projectieve techniek is de thematic apperception test die voor het eerst beschreven werd door Morgan & Murray. Hier wil je een idee krijgen over het leven van een persoon aan de hand van tekeningen. De zin-aanvul test is ook een veel gebruikte projectieve test. De voorstanders van de projectieve technieken zeiden dat je op een indirecte manier een breder beeld krijgt van de persoonlijkheid, en de tegenstanders stellen de betrouwbaarheid en validiteit in vraag.

Kenmerken van deze periode

- Ontwikkeling en gebruik van tests nemen sterk toe.
- Theoretische en methodologische verantwoording ontbreekt vaak

Uitzondering Thurnstone

- Benadrukt belang voor kritische evaluatie van tests (betrouwbaarheid en validiteit).
- Statistiek gaat belangrijke rol spelen in testdiagnostiek.
- Tijdschriften *Psychometrika* en *Educational and Psychological Measurement*.
- Verschillende benadering Europa (fenomenologisch) vs VS (behavioristisch).

Begin WO-II tot heden

In Amerika werden gedurende de oorlogsjaren met succes de meest vooraanstaande psychologen op het terrein van selectie, testontwikkeling en psychiatrisch-medische keuring ingeschakeld bij de gigantische taak waarvoor de keurings- en selectiediensten van de krijgsmacht zich geplaatst zagen. Deze samenwerking leidde tot een enorme toename van het aantal selectie- en diagnostische tests, tot een sterke professionalisering van het selectie- en plaatsingsbeleid en tot een constructieve en kritische bezinning op de psychologische principes van het testonderzoek. Meer dan 9.000.000 personen werden onderzocht met de army General Classification test (AGCT) en hun gegevens vormden een uitstekende bron voor genuanceerde normerings- en valideringsresultaten. In 1947

werd de Educational Testing Service (ETS) opgericht. De bedoeling van deze non-profitorganisatie, met een groot aantal psychologen, onderwijskundigen, statistici en psychometrici in dienst, is tegemoet te komen aan de behoeften van Amerikaanse onderwijs- en opleidingsveld wat betreft de toelating tot en de evaluatie van het onderwijs. Naast ETS is in de V.S. een groot aantal test agencies werkzaam.

Ook op het meer toegepaste terrein van de selectiepsychologie staat, zeker in de jaren na de oorlog, de testtheorie centraal. De theoretische onderbouwingen van de psychometrie en het onderwijskundig meten steeds meer hand in hand gaan. Een volgende belangrijke invloed op de testtheorie is afkomstig uit de schaaltheorie. Ook uit de schaaltheorie voortkomende invloeden, zoals het meten van voorkeuren voor bepaalde stimuli via ontvouwings- en preferentiemodellen, hebben blijvende invloed gehad op de psychometrie. Een andere oorzaak van de versnelde testontwikkeling en testresearch is de ontwikkeling en uitbouw van de testgegevens per computer.

Hoofdstuk 2: Definitie, kenmerken en toepassingen van de test

Metten in de psychologie

Wat is een test?

Een test is een systematische classificatie- of meetprocedure, waarbij het mogelijk is om een uitspraak te doen over een of meer empirisch-theoretisch gefundeerde eigenschappen van de onderzochte of over specifiek niet-testgedrag, door uit te gaan van een objectieve verwerking van reacties van hem/haar, in vergelijking tot die van andere, op een aantal gestandaardiseerde, zorgvuldig gekozen stimuli.

Een verantwoorde en gepubliceerde test bevat testmateriaal en testformulieren. Testmateriaal kan erg divers zijn en is soms geen tastbaar materiaal. Testformulieren dienen voor het verzamelen van materiaal waarop de interpretaties en conclusies gebaseerd zullen worden en in te vullen door de respondent of proefleider.

Een testhandleiding is een grote variatie in uitvoerigheid. Een testhandleiding moet 4 kenmerken hebben:

- Exacte testinstructie: Deze instructie bevat datgene wat betrekking heeft op de gang van zaken tijdens het onderzoek.
- Verwerkingsprocedure: De verwerkingsprocedure bestaat voornamelijk uit de richtlijnen voor de toekenning van numerieke scores aan de antwoorden of de reacties op de opgaven.
- Normtabellen: De score kan worden vergeleken met de prestaties van meer of minder representatieve normgroepen.
- Informatie over wetenschappelijke kwaliteit van de test: Het gaat hierbij om gegevens die een indicatie geven van de betrouwbaarheid van de test, een bespreking van de testbetekenis en de vraag voor welke voorspellingen de test gebruikt kan worden.

De bedoeling van het testonderzoek is het doen van een uitspraak die een voorspelling, classificatie of beschrijving met betrekking tot het onderzochte individu behelst. Een psychologische test is een systematisch onderzoek van gedrag met behulp van speciaal geselecteerde vragen of opgaven, met

de bedoeling inzicht te krijgen in een psychologisch kenmerk van de onderzochte in vergelijking met anderen.

De psychologische test onderscheidt zich van het voorwetenschappelijke oordelen in 6 opzichten:

1. Efficiëntie: Een test is geconstrueerd met de bedoeling een specifiek psychologisch kenmerk te meten. Men is niet afhankelijk van het min of meer toevallig beschikbaar zijn van gelegenheden waarin dat kenmerk zich spontaan manifesteert. De vragen of opgaven zijn zo gekozen dat ze maximaal informatie geven over het kenmerk dat zich wilt meten. Er wordt een situatie gecreëerd waarin bijkomstigheden en toevallige invloeden zoveel mogelijk worden geweerd.
2. Standaardisatie: zelfde condities voor alle personen: noodzakelijke voorwaarde voor vergelijkbaarheid van testprestaties. Er wordt een standaardsituatie gecreeerd zodat men alle geteste personen in zo gelijk mogelijke omstandigheden bezig ziet. Een uitvoerige handleiding is wenselijk. Een perfecte standaardisering is onmogelijk.
Testafname
 - Instructies
 - goed uitgewerkt.
 - proefleider moet zich er zo goed mogelijk aan houden.
 - Testomgeving.
 - opvallende of specifieke omgevingsinvloeden weren (bv. temperatuur, geluid, helderheid van licht, lichtinval, ventilatie, schrijfcondities)
 - storingen voorkomen (afleiders).
 - samenwerken of afschrijven verhinderen
 - proefpersoon
 - vermoeidheid, emotionele opwindings,...
 - (vermeende) verwachtingen proefleider inlossen.
 - sociale wenselijkheid.
 - vertrouwdsheid met specifieke test of testonderzoek in algemeen.
 - Motivatie.
 - Angst.
 - Proefleider.
 - Verwerking
3. Normering: In het normeringsonderzoek is op zijn minst een rangorde vastgesteld van zeer goede tot zeer slechte prestaties, zodat iemands prestatie door middel van een plaatsbepaling in deze rangorde kan worden beoordeeld. Het vaststellen van een rangorde is de eenvoudigste vorm van normering. De benadering met een normaalverdeling doet men doorgaans als de steekproef uit een normaal verdeelde populatie afkomstig is, en men steekproeffouten wil gladstrijken. Het kiezen van nieuwe verdelingskenmerken doet men om een handige schaal te verkrijgen. De beschikbaarheid van normen is meestal een noodzakelijke voorwaarde voor het gebruik van een test. De beschikbaarheid van normen is nodig bij het gebruik van alle soorten tests.
4. Objectiviteit: Objectiviteit wordt opgevat als de onafhankelijkheid van storende invloeden vanuit de persoon van de waarnemer, beoordelaar of interpretator. Voor de testinterpretatie houdt dit in dat men mag aannemen dat het proces van het registreren en verwerken van het testgedrag van de onderzochte tot score, beoordelingscategorie of classificatie, vrij is van

aan de testleider gebonden invloeden. Bij objectiviteit is het resultaat van het onderzoek onafhankelijk van de persoon van degene die gegevens verzamelt of uitwerkt. Dit heeft 2 gevolgen:

- a. Bij het objectieve onderzoek moet het voor de onderzochte niet uitmaken wie de beoordelaar is.
 - b. De objectiviteit impliceert de openheid en reproduceerbaarheid van de test en evaluatieprocedure.
5. Betrouwbaarheid: Met de bepaling van de betrouwbaarheid gaan in de psychologie problemen gepaard. De problemen komen erop neer dat het niet eenvoudig en vaak zelfs onmogelijk is om twee onafhankelijke metingen te verkrijgen.
 6. Validiteit.

Operationaliseren

Meten is belangrijk voor het nagaan van verschillen tussen mensen, fenomenen, verbanden, veranderingen, enz. Eigenlijk om waar te nemen. Meten en gebruik van schalen in de psychologie noodzaak onmeetbare gedragsaspecten meetbaar te maken, dit is een indirecte manier, en spreekt men dus over hypothetische constructen. Empirische constructen of concrete begrippen zijn rechtstreeks waarneembaar. Operationalisering is een directe meting. Meten zijn waarden toekennen aan de verschillende toestanden van een variabele, waarbij de verhouding tussen die waarden een weerspiegeling is van de verhouding in de werkelijkheid.

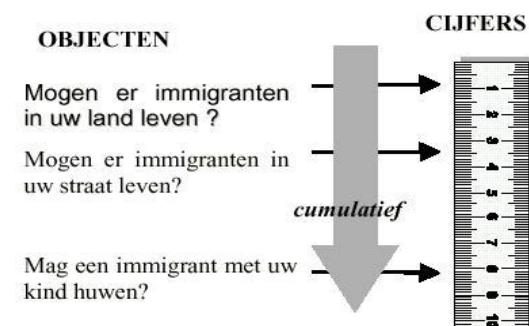
Men kan pas spreken van meten als vanuit het getelde aantal een conclusie volgt over de eigenschap in kwestie. Er zijn stappen om te komen tot een meting:

1. Identificatie van de te meten eigenschap: Hier wordt er vastgesteld wat de theorie is met betrekking tot de te meten eigenschap, of worden de rivaliserende theorieën inzake deze eigenschap in kaart gebracht. Het probleem bij de identificatie van de eigenschap en de keuze van theoretische uitgangspunten voor de testconstructie is dat met betrekking tot vele psychologische eigenschappen de theorievorming maar matig is ontwikkeld, terwijl voor andere eigenschappen geldt dat de status van deze theorieën onduidelijk of dubieus is. Het grootste probleem vormt de categorie van eigenschappen met een theoretische onduidelijke of dubieuze status.
2. Operationalisering van de eigenschap: Met operationalisering van een eigenschap wordt bedoeld de specificatie van de operaties die nodig zijn om de eigenschap te meten. De operationalisering vereist ten eerste dat een domein van gedragingen wordt gedefinieerd die typisch zijn voor de bedoelde eigenschap. Het gevolg is dat het gedragsdomein moeilijk af te bakenen is. De tweede stap is het hierop afstemmen van een domein van stimuli die aan proefpersonen kunnen worden gepresenteerd om reacties uit te lokken die een indicatie geven van de bedoelde eigenschap.
3. Onderzoek en kwantificering van reacties: De reacties van personen op items geven informatie over de eigenschap waarover de test wordt geacht uitsluitsel te geven. We gaan ervan uit dat bij een item het ene antwoord een hoger niveau aangeeft op een achterliggende eigenschap dan het andere antwoord. De kwantificering bestaat erin dat aan de reacties in de hogere categorie een hogere numerieke waarde wordt toegekend.

4. Terugkoppeling naar de theorie: De uitkomsten van de statistische analyses hebben repercussies voor de theorie over de te meten eigenschap. Blijkt na onderzoek de theorie te kloppen, dan betekent dat er een aanvullende evidentie is gevonden voor de theorie van de achterliggende eigenschap. Maar ook een niet-succesvolle of anders dan verwacht verlopen testconstructie dragen bij aan theorievorming. Er kan een aanleiding gevormd worden om de theorie te herzien of kiezen voor een van de rivaliserende theorieën over een eigenschap. Representational measurement gaat ervan uit dat er alleen sprake kan zijn van meting, indien de formele relaties tussen meetwaarden overeenkomen met de relaties die er in de empirie bestaan tussen objecten waarover die meetwaarden iets zeggen. Een dergelijke directe waarneembaarheid van eigenschappen en relaties wordt in de psychologie zelden aangetroffen. Representational measurement en het meten van psychologische, hypothetische eigenschappen laten zich niet goed verenigen, toch is representational measurement van groot belang voor het denken over meten in de psychologie.

Antwoordschalen, antwoordvormen

Inschalen of schaleren is het toekennen van cijfers aan objecten of aan waarnemingen, betreffende objecten volgens een regel. Objecten is het laagste niveau van je test, bv. stellingen of vragen. Een antwoordschaal of responsschaal is de manier waarop je de antwoorden op beweringen verzamelt. Een (meet)schaal is een verzameling van mogelijke categorieën of scores van een variabele scores voor de verschillende objecten die we bekomen via een welbepaalde regel.



Meetschalen hebben als doel om een bepaald psychologisch construct in meting te brengen. Je moet je baseren op de antwoordschaal dat je gebruikt. De meetschaal of schaalscore wordt verkregen door de som van alle vragen of welbepaalde items op te tellen.

Psychologische tests en vragenlijsten bestaan uit basiselementen die items worden genoemd. Door middel van items wordt respondenten gevraagd activiteiten te verrichten, op basis waarvan een psychologische eigenschap wordt beoordeeld:

- Theoretische opdrachten: Antwoorden op opdrachten zijn goed of fout, of soms gedeeltelijk goed of fout.
- Stellingen: De antwoorden zijn niet goed of fout, maar de positie van de respondent op de schaal van de veronderstelde eigenschap aangeven.
- Vragen: De antwoorden van respondenten op vragen bieden informatie over dergelijk eigenschappen. Antwoorden zijn hier niet goed of fout, maar geven een standpunt of een houding aan.
- Praktijkproeven: Bij dit soort opdrachten kunnen reacties diverse vormen aannemen. Wat hier aan activiteit gevraagd wordt is meestal een juist zo 'vrij' mogelijke reactie.

De reacties van respondenten op items kunnen grofweg in twee categorieën verdeeld worden. Ten eerste is er de mogelijkheid dat de respondent geheel zelf bepaalt hoe hij reageert (open vragen).

Ten tweede is er de mogelijkheid dat een respondent kiest uit een beperkt aantal door de testconstructeur geformuleerde reacties (gesloten vragen).

De antwoorden op open vragen bieden vaak meer informatie dan de gesloten vragen. Maar het beoordelen en categoriseren van de antwoorden op open vragen vraagt veel tijd voor de onderzoeker, maar ook voor de respondent. Ten tweede begrijpen respondenten niet altijd wat er met een vraag bedoeld wordt en soms kunnen antwoorden irrelevant zijn. Ten derde zijn niet alle respondenten even goed in staat om hun gedachten op papier te zetten.

Bij gesloten vragen gaat ten eerste de beantwoording door de respondent en de categorisering van de antwoorden door de onderzoeker zeer snel. Ten tweede krijgt de respondent geen gelegenheid om op voor het onderzoek irrelevante zaken in te gaan en is de kans op het verkeerd begrijpen van de stelling kleiner. Ten derde speelt de interpretatie van het antwoord door de onderzoeker geen rol.

Omdat bij open vragen altijd de subjectiviteit van de beoordeling een rol speelt en deze subjectiviteit ook bij een goede voorbereiding door de onderzoeker nooit geheel kan worden uitgesloten, moet in de fase van de testconstructie de categorisering van de reacties door ten minste twee beoordelaars geschieden.

Het eerste nadeel van meerkeuzevragen is dat het lastig is om ze te bedenken. Ten tweede is het inschatten van de moeilijkheidsgraad van dit type items lastig. Het praktische voordeel van meerkeuzevragen zit in de geringe tijd die besteed hoeft te worden aan correctie en toekennen van scores. Het tweede bezwaar tegen meerkeuze is de kans dat respondenten door louter te raden vele goede antwoorden kunnen krijgen. Ten derde is dat er met meerkeuzevragen enkel oppervlakkige cognitieve vaardigheden mee zouden kunnen gemeten worden.

In prestatieniveau tests wordt vaak gebruik gemaakt van meerkeuze-items. Doorgaans is een oplossing goed, en de anderen fout. De uitspraak wordt de stam genoemd, en de antwoordmogelijkheden worden onderscheiden in de gesleutelde respons (het correcte antwoord) en de afleiders (de foute antwoorden). Je hebt ook nog de keuze tests. De eenvoudigste vorm van kiezen is de tweekeuze vorm (bv.: Ja/neen). Varianten op de meerkeuze-items met meer dan 2 mogelijkheden zijn:

- (in)correcte antwoordvorm: Slechts een van de antwoorden is het juiste of het onjuiste.
- Meest/minst juiste antwoordvorm: De respondent moet uit 4 meer of minder plausibele antwoorden het meest juiste of het minst juiste kiezen.
- Verschillende antwoordenvorm: Uit een aantal antwoordmogelijkheden dient de respondent een of meer antwoorden te kiezen die een juist antwoord zijn op de gestelde vraag.
- Complexe vormen.

Bij rangschikking moet de respondent de keuzemogelijkheden plaatsen in volgorde van juistheid, toepasbaarheid of voorkeur. Het voordeel van rangschikking is dat veel informatie wordt verkregen. Als laatste heb je dan nog toeschrijving, ook wel matching genoemd. uit een rij namen, gebeurtenissen of feiten en een rij eigenschappen, oorzaken of jaartallen moet telkens de juiste combinatie gemaakt worden. Het voordeel van deze methode is dat door de compacte vraagvorm veel informatie door middel van een opgave wordt verzameld.

In tests voor gedragswijze is een item vaak een uitspraak waarbij de respondent, door op een gegeven schaal, een kruisje te zetten, kan aangeven in welke mate hij het eens is met de uitspraak. De presentatievorm van dergelijke items kan sterk variëren. Enkele belangrijke aanwijzingen en overwegingen bij dergelijke keuzes zijn:

- Zorg ervoor dat de prestatievorm en instructie zodanig zijn dat respondent niet in verleiding komt om een kruisje tussen 2 mogelijkheden te zetten.
- Het meest bekend zijn de items met de 5 antwoordenmogelijkheden, ook wel likert schaal genoemd.
- Soms wordt weleens gedacht dat een zeer groot aantal keuzemogelijkheden beter is dan een kleiner aantal, omdat de respondent dan een genuanceerde mening kan geven.
- Een andere keuze betreft de verbale aanduiding van opties.

De reacties van respondenten zijn in eerste instantie kwalitatief van aard. Met deze kwalitatieve reacties is het moeilijk om systematisch onderzoek te doen naar de kwaliteit van de items en van de test of vragenlijst als geheel. Bovendien dienen aan personen meetwaarden te worden toegekend en op basis van de kwalitatieve reacties is ook dat niet goed te realiseren. Om dit laatste te kunnen bereiken, dienen de kwalitatieve reacties op de items in getallen te worden omgezet.

Dimensionaliteit

Dimensionaliteit heeft te maken met het feit of een construct nog relevante onderliggende deelaspecten of dimensies bevat.

Dimensionaliteit van meetschalen

Schalen kunnen bestaan uit een of meerdere dimensies:

1. Uni-dimensionele schalen: Vb. Het meten van hoogte. Hier wordt er maar een enkel algemeen kenmerk gemeten, de hoogte.
2. Twee-dimensionele schalen: Vb. Intelligentie. Hierbij meet je de verbale en de performale schaal. Men heeft dus 2 cijfers nodig.
3. Multi-dimensionele schalen: Vb. Causale attributietheorie, nl. controleerbaar, stabiel en Locus.

Likert schaal

Een likertschaal bestaat uit een reeks items die uitspraken bevatten. De respondent moet bij elk item aangeven in welke mate hij/zij akkoord is met de uitspraak. Deze schaalscore wordt dan geïnterpreteerd als meetwaarde voor het bestudeerde construct. Voor negatief geformuleerde items draaien we de scoring om.

Dit is een voorbeeld van een meetschaaltipe. Een bepaalde regel over het verband tussen het te meten onderliggend construct en de manier waarop je e antwoorden op een schaal worden gegeven. De schaalscore is de som van cijfers die horen bij de antwoorden. Schaalscore is de meetwaarde voor het construct. De regel bij de likert schaal geldt: Voor elk item neemt de kans dat de mate waarin iemand ermee akkoord gaat toe of af, afhankelijk van de positie van die persoon op de dimensie van het construct.

Stappen bij constructie van unidimensionele Likert schaal:

1. Definitie van unidimensionele construct.
2. Groot aantal items opstellen die aansluiten bij definitie.
3. Items voorleggen aan proefgroep (< populatie waarop men schaal wil toepassen).
4. Analyse itemscores en schaalscores van proefgroep.

Meetniveau

Je hebt 4 meetniveaus:

1. Nominale schaal: Als we personen of objecten volgens een bepaald principe verdelen in groepen en we kennen aan doe groepen een getal toe ter identificatie, dan is dit een vorm van meten. Alleen al categorisering en naamgeving is volgens deze opvatting een vorm van meten. Dit eerste niveau van meten vind plaats op de nominale schaal. Aangezien getallen hier alleen dienen om categorieën of objecten te onderscheiden, maakt het in principe niet uit welke getallen men kiest.
2. Ordinale schaal: Een ordinale schaal wordt gekenmerkt door het begrip rangorde. Getallen worden hier gebruikt om categorieën of objecten te ordenen. De keuze van getallen is derhalve vrij zolang zij de juiste ordening blijven weerspiegelen.
3. Intervalschaal: Op de intervalschaal zijn niet alleen de verschillende posities geordend, maar hebben de afstanden tussen de posities een betekenis.
4. Verhoudingsschaal of ratio schaal: Deze schaal heeft eenheden waarmee afstanden aangegeven kunnen worden, maar het heeft ook een absoluut nulpunt.

Slechts weinig metingen in de psychologie vinden plaats op de verhoudingsschaal. Overigens vergen de meeste toepassingen van de psychologie hoogstens een intervalschaal en is een ordinale schaal al vaak voldoende.

Scoring- en transformatiemeetwaarden

Ruwe scores en transformatiemeetwaarden

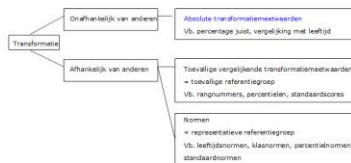
De ruwe score is de basisscore waarin een testresultaat in eerste instantie wordt uitgedrukt. Het cijfer op zich zou een idee moeten geven van de betekenis van de bepaalde testscore. Een verzamelnaam voor dergelijke cijfers is transformatiemeetwaarden. Men transformeert of past de ruwe scores aan zodat het resultaat op een test interpreteerbaar wordt. Absolute transformeerwaarden zijn cijfers die onafhankelijk zijn van de prestaties van anderen op die test. Vergelijkende transformatiewaarden zijn cijfers die tot stand komen door een vergelijking te maken tussen de ruwe score van een persoon met de scores behaald door anderen. De vergelijkende transformatiemeetwaarden zijn nog eens in te delen op basis van toevallige of representatieve vergelijking:

- Toevallige vergelijkende transformatiemeetwaarden zijn cijfers die een vergelijking uitdrukken tussen de score van een persoon en de scores behaald door een toevallige groep anderen.
- Normen daarentegen zijn transformatiemeetwaarden die testcores uitdrukken in vergelijking tot scores behaald door een representatieve groep van vergelijkbare anderen.

2.6 SCORING EN TRANSFORMATIEMEETWAARDEN

2.6.2 ABSOLUTE TRANSFORMATIEMEETWAARDEN

Transformatiemeetwaarden



Absolute transformatiemeetwaarden

De testuitslag wordt eerst uitgedrukt in een basisscore, die bekend staat als ruwe score. De ruwe score is de eerste testscore die wordt berekend en die nog geen verdere bewerking heeft ondergaan. De ruwe score heeft op zichzelf meestal weinig betekenis. Er kunnen 3 soorten bewerkte scores geïllustreerd. Het eerste type is gebaseerd op vergelijking met een absolute standaard. Het tweede type is gebaseerd op een deling. En een derde type score is gebaseerd op de relatieve positie in een referentiegroep.

Percentage juist

Onder een norm verstaan we een referentiekader voor de evaluatie van de ruwe scores dat gebaseerd is op de kenmerken van de verdeling van de ruwe scores in een populatie. Deze kenmerken worden geschat op basis van een representatieve steekproef. Het gebruik van de normen voor de beoordeling van een concrete testscore houdt de vergelijking in van deze score met de scoreverdeling van de referentiepopulatie. Niet altijd is het nodig of zelfs gewenst normen te berekenen of te gebruiken. Er zijn gevallen waar men zich kan beperken tot ruwe scores.

De overgang van een vergelijking binnen een onderzochte groep naar het gebruik van normen is vloeiend. Naarmate de onderzochte groep groter is en meer de kenmerken van een relevante populatie weerspiegelt, krijgen de vergelijkingen meer het karakter van vergelijkingen met een norm. Hierbij zijn 2 zaken belangrijk:

1. Het is van belang bij normeringsonderzoek de specifieke kenmerken van de onderzochte normgroep te vermelden.
2. Men dient normen niet al te absoluut te zien.

Vergelijking met leeftijd

Verhoudingsnormen hebben vooral een historische betekenis. Bij deze vorm van scorebewerking worden de testcores gedeeld door een andere variabele en daardoor onafhankelijk gemaakt van die variabele.

Vergelijkende transformatiemeetwaarden

2.6 SCORING EN TRANSFORMATIEMEETWAARDEN

2.6.3 VERGELIJKENDE TRANSFORMATIEMEETWAARDEN



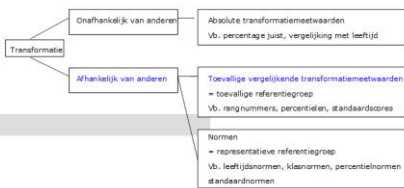
Voordeel 'val' transformatiemeetwaarden
Impact van moeilijkheid van test vermindert voor de interpretatie van een score

Wanneer men een toevallige referentiegroep gebruikt voor de interpretatie van een testscore, dan zal men voorzichtig moeten zijn, omdat de interpretatie sterk afhankelijk zal zijn van de samenstelling van die toevallige referentiegroep. Wanneer men een representatieve vergelijkingsgroep gebruikt zal men de testscore van een persoon niet enkel kunnen interpreteren in termen van vergelijking met de referentiegroep, maar ook ten opzichte van de populatie. De transformatiemeetwaarden die men dan bekomt bij een

representatieve referentiegroep, noemt men normen.

2.6 SCORING EN TRANSFORMATIEMEETWAARDEN 2.6.3 VERGELIJKENDE TRANSFORMATIEMEETWAARDEN

2.6.3.1 Toevallig vgl. transformatiemeetwaarden



Toevallige vergelijkende transformatiemeetwaarden

Rangnummers

De eenvoudigste vorm van vergelijking van testprestaties tussen individuen is de rangordening. De bewerkte score is het rangnummer. De nadelen zijn dat he de prestaties nog steeds moeilijk kan inschatten en de interpretatie is afhankelijk van de kwaliteit of aard van de groep.

Fractielen: Percentielen, decielen, quartielen

Fractielen verwijst naar het indelen van de verdeling van verkregen scores in gelijke fracties volgens een bepaalde regel. Kennis van de groepsgrootte is niet vereist bij percentielscores, een zeer populaire vorm van scorebewerking. Percentielen zijn die 99 punten, die een frequentieverdeling verdelen in 100 groepen van gelijke grootte. De percentielscore wordt berekend door middel van lineaire interpolatie. Percentielscores worden veel gebruikt. De redenen zijn de eenvoudige en snelle berekening, de gemakkelijke toepasbaarheid en de inzichtelijkheid, ook voor de niet-deskundige. Ze hebben echter ook nadelen, omdat de schaal van percentielscores ordinaal is. Ten eerste mogen met percentielscores geen gemiddelden en varianties worden berekend. Ten tweede heeft het geen zin om de frequentieverdelingen van percentiele en ruwe scores te vergelijken.

Een ruwere fractielform is de decielscore. Hierbij wordt de verdeling van ruwe scores ingedeeld in 10 zones waarbinnen telkens 10% van de observaties liggen. Het voordeel van fractielen is dat ze antwoord bieden aan het probleem van de grootte van de referentiegroep en dat ze gemakkelijk toepasbaar en inzichtelijk zijn. De nadelen van fractielen zijn de kwaliteit van de referentiegroep en de afstanden tussen de ruwe scores die niet in de fractielen blijven.

Standaardscores: Z score, T score, deviatie IQ, Stanines, C score

2.6.3.1 TOEVALLIGE VGL. TRANSFORMATIEMEETWAARDEN STANDAARDScores Z SCORES

lineair	genormaliseerd
Hoeveel standaardafwijkingen ligt de ruwe score boven of onder het gemiddelde in de WAARGENOMEN verdeling?	Waargenomen verdeling wordt eerst omgezet naar een normaalverdeling. Hoeveel standaardafwijkingen ligt de score boven of onder het gemiddelde in de NORMAALverdeling?
$Z_{X_i} = \frac{X_i - \bar{X}}{s_x}$	

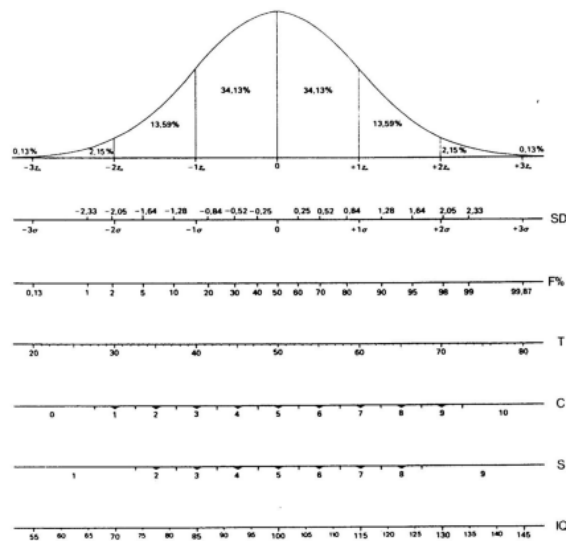
Ruwe scores die onder het gemiddelde liggen corresponderen met negatieve standaardscores en ruwe scores die boven het gemiddelde liggen corresponderen met positieven standaardscores. Standaardscores behouden, op het gemiddelde en standaarddeviatie na, dezelfde verdeelingskenmerken als de ruwe scores, X. Genormaliseerde standaardscores worden verkregen uit niet-lineaire transformaties die de verdeling van X zodanig vervormen dat wel een normaalverdeling ontstaat. De constructie van deze genormaliseerde standaardscores vindt plaats conform de

hiergenoemde gedachte. Men zoekt uit welke ruwe score juist die percentages markeren, die behoren bij standaardeenheden in een normaalverdeling. Lineaire transformatie: standaardscores behouden dezelfde verdeelingskenmerken als ruwe scores: Z-scores zijn dus niet automatisch normaal verdeeld! Bij omzetting van ruwe scores naar Z-scores wordt de verdeling verschoven (gemiddelde wordt van elke score afgetrokken), wordt de afstand tussen de scores veranderd (elke score wordt gedeeld door de standaarddeviatie), maar blijven andere kenmerken (bv. scheefheid) gelijk.

Genormaliseerde standaardscores zijn een niet-lineaire transformatie: De verdeling van scores wordt zodanig vervormd dat er een normale verdeling ontstaat. Tot de categorie van de genormaliseerde standaardscores behoren de volgende meest voorkomende vormen:

- T-scores: Bij omzetting van ruwe scores X naar T-scores wordt een gemiddelde aangenomen van 50 en een spreiding van 10.
- Stanines.
- Deviatie-IQ: Het gaat hier om standaardnormen met een gemiddelde van 100.
- C-score: Hierbij wordt de volledige scoreverdeling in 11 klassen van 0.5 standaarddeviaties ingedeeld.

Verband tussen transformatiemeetwaarden



Normen

Een norm is de referentiekader voor de evaluatie van de ruwe scores dat is gebaseerd op de kenmerken van de verdeling van de ruwe scores in een populatie. Deze kenmerken worden geschat op basis van een representatieve steekproef. Het is belangrijk om de specifieke kenmerken van de normgroep te vermelden in handleiding. Maar: normen zijn niet absoluut - tijdgevoelig; geregelde revisie noodzakelijk, bij vertaling van een test kan men niet zomaar de normgegevens uit het oorspronkelijke taalgebied overnemen. Een norm zijn transformatiemeetwaarden die werden bekomen voor een representatieve referentiegroep. Normeren is het vastleggen van de normen van een test door test af te nemen bij een representatieve steekproef. Ruwe scores vergelijken met norm, onafhankelijk van prestatie in toevallige groep. Normen zijn beschikbaar voor diverse deelpopulaties. Ruwe subtestscores worden omgezet naar stanines.

Profielen

Norm-, criterium- en communicatiegeoriënteerde instrumenten

Bij de vergelijking van een testscore met een absolute standaard wordt de testprestatie beoordeeld zonder die van anderen hierbij te betrekken. De vraag is dan hoe goed de prestatie is in vergelijking met een absolute maatstaf of norm, en niet wat de prestatie is in vergelijking met de klas, leeftijdsgenoten of Nederlandse populatie. Deze maatstaf kan subjectief en arbitrair zijn. De Engelse term is criterion-referenced measurement. Het vergelijken van een testprestatie met een normverdeling heet norm-referenced measurement. De Nederlandse termen zijn absoluut en normatief meten.

Bij absolute normering is een analyse van de doelen van het proces dat men bij de onderzochte wil evalueren – een leerproces, een aanpassingsproces of een training – essentieel, evenals de vraag of de mate waarin die doelen zijn verwezenlijkt op een betrouwbare manier te meten valt. Is dit laatste niet het geval, dan krijgt absoluut meten een sterk arbitrair karakter en is de rechtvaardiging van deze vorm van meten discutabel.

- Normgeoriënteerde instrumenten of 'Norm referenced' measurements
 - Normgroep of referentiegroep.
 - Vergelijkende transformatiemaatwaarden.
 - Gebruik binnen diagnostisch proces.
 - screening grote groepen.
 - bepalen van de aanwezigheid en ernst van een afwijking of achterstand.
- Criteriumgeoriënteerde instrumenten = 'Criterion referenced' measurements of (criterium)toetsen.
 - Vergelijking tussen het gedrag van de onderzochte persoon met gewenst gedrag; vastgelegd doel.
 - Gebruik binnen onderwijscontext en diagnostisch proces.
 - Plannen van interventie.
- Communicatie georiënteerde instrumenten = Communication-referenced measurements
 - Doel: Het ontdekken van het individueel beleven en functioneren van de onderzochte persoon.
 - Geen vergelijking met groep of expliciet criterium.

Aard van het gemeten gedrag: testindeling

Testgedrag

Vrijwel alle testhandboeken en testoverzichten gaan uit van de indeling van tests op basis van het testgedrag. Het belangrijkste onderscheid dat uitgaat van het geteste gedrag is dat tussen tests voor prestatieniveau en tests voorgedragwijze. Bij tests voor prestatieniveau wordt de onderzochte persoon een maximale prestatie gevraagd. Testvragen worden goed of fout beantwoord en er is een duidelijke norm wat goed of fout is. Bij tests voor gedragwijze daarentegen is niet een vooraf bekenden goed/fout-sleutel beschikbaar. Het gaat vooral om de vraag hoe iemand iets doet of op welke wijze de prestatie of reactie plaatsvindt. Het gedrag of de prestatie zijn niet van belang in hun verhouding tot een maximale prestatie.



Tests voor prestatieniveau

Enkelvoudige algemene niveautests

De grootste en belangrijkste groep tests wordt gevormd door de enkelvoudige algemene niveautests, die leiden tot een enkele indicatie van de intelligentie. Binnen deze categorie kunnen 3 deelcategorieën worden onderscheiden:

1. De individuele ontwikkelingstests: Bij ontwikkelingstests wordt de intelligentiescore meestal vergeleken met de testprestaties van leeftijdsgenoten. Ook noemen we in deze categorie de intelligentietest voor visueel gehandicapte

kinderen. Deze test is gebaseerd op het model van Thurstone met 7 intelligentiefactoren. De batterij bestaat uit diverse verbale en haptische deelttests, waarvan enkele zijn ontleend aan de WISC en RAKIT. Evenals bij de WISC en RAKIT worden er scores op de deelttests bepaald, die vervolgens worden gecombineerd tot een algemene intelligentiescore.

2. De individuele intelligentietests voor volwassenen: De individuele intelligentietest voor volwassenen is gericht op het vaststellen van het intelligentieniveau wanneer de intelligentie min of meer volgroeid kan worden beschouwd.
3. De algemene collectieve intelligentietests: De collectieve intelligentietests kunnen groepsgewijs worden afgenomen wanneer het testen grootschalig is en efficiëntie belangrijk. Soms worden items die verschillende eigenschappen meten dooreen gemengd, en als zodanig voorgelegd, soms worden ze op a priori of empirische gronden gegroepeerd. Het gaat bij deze tests om een indicatie van het algemene intelligentieniveau. De meeste collectieve algemene intelligentietests zijn breed en gevarieerd samengesteld. Verder wordt er de collectieve niet-verbale tests, die wat betreft het intelligentiedomein beperkter zijn dan de collectieve algemene intelligentietests, omdat er geen gebruik gemaakt wordt van verbaal materiaal, maar pantomimisch gegeven.

Veelvoudige algemene niveautests

Bij veelvoudige algemene niveautests gaat het juist wel om een nadere differentiatie en niet alleen, of soms helemaal niet, om het algemene intelligentieniveau. Tests voor differentieel gebruik moeten aan andere voorwaarden voldoen dan de enkelvoudige algemene niveautests. De belangrijkste voorwaarde voor de veelvoudige algemene niveautests is dat de deeltests afzonderlijk gezien voldoende betrouwbaar en tevens onderling voldoende onafhankelijk zijn. Er zijn twee typen van veelvoudige algemene niveautests:

1. Testbatterijen voor intelligentiefactoren: Het doel is hier dat elke deeltest een homogene dimensie van intelligentie meet. Het meest uitgewerkte alternatief voor Thurstone's factoren is het systeem van intelligentiefactoren volgens Guilford. Hierin is een grote hoeveelheid, merendeels experimentele, tests opgenomen voor een grote variëteit van intellectuele taken. Door herhaalde analyse van deze en soortgelijke tests leidde Guilford een factoranalytische structuur van de intelligentie af, die aanvankelijk een empirisch karakter had, maar allengs een meer theoretisch karakter kreeg.
2. Testbatterijen voor geschiktheden: Hier wordt er meer verwezen naar een maatschappelijke taak of een schooltaak en minder naar de structurele opbouw van intelligentie, waarin men van al of niet zuivere abilities uitgaat. De testbatterij voor geschiktheden pretendeert de vermogens te meten die iemand in staat stellen zich voor een maatschappelijke taak of schooleisen te bekwamen. Soms zijn de geschiktheidsfactoren samengesteld uit diverse intelligentiefactoren.

Speciale niveautests

Speciale niveautests zijn toegespitst op een bepaald segment van de begaafdheid. We onderscheiden 3 tests in dit gebied:

1. Tests voor speciale intelligentiefactoren: Bij tests voor speciale intelligentiefactoren gaat het om tests die uitsluitend eigenschappen als ruimtelijk inzicht, woordkennis of numerieke vaardigheid meten. Hiervan is een groot aantal beschikbaar. Een wat aparte plaats wordt ingenomen door de redeneertests of abstractietests. Guilford heeft in zijn intelligentiemodel tevens wat meer ongewone, maar praktisch gezien wel degelijk belangrijke intelligentie-aspecten binnen het gezichtsveld van de testpsychologie gebracht. Een andere ontwikkeling is die in de behavioral dimensie van Guilford, voorheen wel de sociale intelligentie genoemd. Onderzoek heeft echter laten zien dat sociale intelligentie in het algemeen slecht te onderscheiden is van andere intelligentiefactoren.
2. Tests voor speciale geschiktheden: Onder de tests voor speciale geschiktheden vallen allereerst de tests voor leergeschiktheden. Hieronder kunnen worden gerekend de tests die zich richten op factoren die men leercondities zou kunnen noemen. Voorts vallen hieronder de tests die gericht zijn op specifieke leergeschiktheden, zoals tests voor technische aanleg.
3. Tests voor speciale niet-intelligentiefactoren: Een laatste groep tests die onder de speciale niveautests valt, wordt gevormd door de tests voor speciale niet-intelligentiefactoren. Het

gaat hier om capaciteiten of vaardigheden waarvan de meting met het oog op een maatschappelijke functie of in het kader van de klinische diagnostiek van belang kan zijn.

- ➔ Motoriektests.
- ➔ Artistieke tests.
- ➔ Sensorische tests.

Vorderingstests

Met vorderingstests tracht men te meten in hoeverre de onderzochte het doel van een opleiding heeft bereikt. Een vorderingstest is niet hetzelfde als een kennistest, omdat het verkrijgen van kennis lang niet uitsluitend, en soms zelfs helemaal niet, het doel van een opleiding is. In een goede kennistest zijn de vragen op een zeer zorgvuldige wijze geselecteerd, geformuleerd en in voorafgaand onderzoek op hun psychometrische kenmerken en hun bijdrage tot het doel van het onderzoek onderzocht. Als in de opleiding ook andere doeleinden dan kennisvermeerdering worden nagestreefd, zal de vorderingstest andere eigenschappen hebben.



Tests voor gedragswijze

Het meten van gedragswijze kent in vergelijking met het meten van prestatieniveaus drie specifieke problemen:

1. Tests voor gedragswijze zijn moeilijk te generaliseren van een kunstmatige testsituatie naar de praktijk of een maatschappelijk functioneren.
2. Er zijn weinig objectieve gegevens die als criterium kunnen dienen bij tests voor gedragswijze.
3. Persoonlijkheidstreken zijn vaak minder stabiel, minder over situaties generaliseerbaar en

Observatietests

In tegenstelling tot andere tests voor gedragswijze komt bij observatietests de informatie steeds van een onafhankelijke beoordelaar die de onderzochte observeert. De gegevens zijn dus indirect, en zij vormen de basis van het interpretatieproces van de psycholoog.

1. **Individuele observatietests:** Bij de individuele observatietests wordt slechts een respondent geobserveerd. Deze persoon wordt onderworpen aan een opdracht die zoveel mogelijk uiteenlopende gedragingen en reacties kan oproepen. De rol van de observator kan variëren van een volledige betrokkenheid bij en deelname aan het proces of het gesprek tot die van een neutrale toeschouwer, waarbij de geobserveerde zich overigens meestal wel degelijk bewust is van de aanwezigheid van de observerende testleider.
2. **Groepsobservatietests:** Bij groepsobservaties gaat het om de beoordeling van een groep mensen, die gezamenlijk een opdracht verrichten. In eerste instantie denkt men bij de groepsobservatietests aan een groep onderzochte personen die een taak uitvoert en een observator die de gedragingen en prestaties registreert en interpreteert. Een andere wijze van gegevensverzameling is die waarbij de deelnemers elkaar beoordelen en de psycholoog deze beoordelingen nader verwerkt.

Somato-fysiologische methoden

Tot de soma-fysiologische methoden behoren alle procedures die door de meting van lichamelijke kenmerken of processen een indicatie trachten te geven van de psychologische kwaliteiten van de onderzochte. Een voorwaarde voor de juistheid van deze methodiek is dat er een samenhang bestaat

tussen fysieke of fysiologische aspecten en gedragskenmerken. Er zijn 2 soorten somato-fysiologische methoden te onderscheiden:

1. Morfologische methoden: Zuivere somatische kenmerken worden onderzocht als mogelijke indicaties voor persoonlijkheidskenmerken.
2. Fysiologische methoden: Deze kunnen nog eens onderverdeeld worden in verschillende methoden:
 - a. Biochemische indices: voor de registratie van hormonale en andere biochemische processen.
 - b. Elektro-encefalogram: voor de meting van spanningsverschillen en –schommelingen in het verloop hiervan in de hersenen.
 - c. Andere 'brain scanners': voor de bestudering van vooral disfunctionerende hersenfuncties.
 - d. Elektrocardiogram: voor de psychofysiologie belangrijke frequentie en regelmaat in de hartslag.
 - e. Metingen van bloeddruk en bloedvolume.
 - f. Registratie van oogbewegingen, oogposities en pupildiameter.
 - g. Meting van elektrodermale verschijnselen: Vb. Skin potential reflex.

Deze fysiologische maten hebben relaties met een veelheid van psychologische variabelen die voor het persoonlijkheidsonderzoek van belang zijn.

Zelfbeoordelingen

Zelfbeoordelingen hebben meestal de vorm van vragenlijsten. Zelfbeoordelingen zijn lange tijd gewantrouwd omdat de respondent onvoldoende objectief zou zijn om zijn eigen functioneren te beoordelen. Er is echter ook een ander gebruik van zelfbeoordelingen mogelijk, dat is gebaseerd op empirisch vergelijkingsonderzoek en dat doorgaans leidt tot meer valide conclusies. Types tests hiervan zijn:

1. Interesstetests: Interesstetests hebben betrekking op het gebied van de interesses en dienen om de beslissing voor school of beroepskeuze te ondersteunen.
2. Waarde- en attitudetests: De attitudetest die bekend is als California-F-Scale wordt gebruikt om na te gaan in hoeverre iemand dogmatisch conservatief en bevooroordeeld staat tegenover diverse politieke, culturele en sociale verschijnselen.
3. Biografische vragenlijsten: zijn gesystematiseerde, schriftelijk afgenomen anamneses. De te voorspellen criteria zijn vaak nogal specifiek, terwijl het aantal biografische gegevens meestal groot is en bovendien een complexe structuur heeft.
4. Persoonlijkheidsvragenlijsten: zijn gericht op een ruime keuze en een rijke variëteit aan aspecten van de persoonlijkheid. Persoonlijkheidsvragenlijsten kunnen op drie manieren worden geconstrueerd:
 - ➔ De vragenlijsten kunnen zuiver empirisch zijn samengesteld op basis van hun relatie met een relevant geacht criterium.
 - ➔ Vragenlijsten kunnen ontstaan uit een grote hoeveelheid persoonlijkheidsitems die met behulp van statistische methoden zijn gegroepeerd.
 - ➔ Vragenlijsten kunnen primair gebaseerd zijn op een theorie omtrent de persoonlijkheid of het temperament.

Voor zelfbeoordeling waarbij de persoonlijkheid in engere zin wordt gemeten, zijn er ook methoden die geen gebruik maken van vragenlijsten. Hier zijn 3 voorbeelden:

1. Osgood-schalen = semantische differentiaal: Hierbij wordt de onderzochte gevraagd een serie eigenschappen, gebeurtenissen of objecten te beoordelen op een aantal bipolaire dimensies.
2. Role Construct Repertoire (Kelly): Er wordt uitgegaan van de veronderstelling dat de ontwikkeling van de persoonlijkheid bestaat uit de ontwikkeling van een serie constructs en dat ieder mens een unieke serie constructs heeft.

3. Q-technique of Q-sort: Men vraagt de onderzochte een serie uitspraken te sorteren, in bijvoorbeeld 9 stapeltjes, van gelijke grootte of in aantallen die de normaalverdeling volgen.

Kwalitatieve prestatietests

Bij een kwalitatieve prestatietest krijgt de onderzochte een opdracht waarvan de vervulling een prestatie lijkt. Het gaat echter niet om de vraag, maar om de wijze waarop deze wordt uitgevoerd. Er zijn 2 soorten kwalitatieve prestatietests:

1. Niveautesen voor gedragswijze: Deze tests kunnen nog eens onderverdeeld worden in kleinere tests:
 - a. Experimentele tests.
 - b. Motoriektests.
 - c. Intelligentietests voor klinisch of diagnostisch gebruik.
 - d. Karaktertests.
 - e. Cognitieve of perceptuele stijlen.
2. Projectietests: Ook deze kunnen onderverdeeld worden in kleinere tests:
 - a. Perceptietests: Vb. Rorschach.
 - b. Interpretatietests: Vb. Thematic apperception tests.
 - c. Expressietests: Vb. Boomtest

Individuele versus groepstests

Bij de individuele test is er sprake van een individuele relatie tussen de testleider en de onderzochte. De vragen of opdrachten worden stuk voor stuk door de testleider geformuleerd en ook de reacties en prestaties worden meestal door de testleider vastgelegd. Bij de groepstest bestaat de activiteit van de testleider slechts uit het geven van de collectieve instructie voor een hele groep proefpersonen tegelijk. In Europa is de individuele wijze van testen nogal populair. Afgezien van de geringe efficiëntie heeft individueel testen zeker voordelen:

- Er is de mogelijkheid de onderzochte te stimuleren als de indruk ontstaat dat hij zijn best niet doet.
- Men kan beter nagaan of de onderzochte de opgave heeft begrepen.

Meestal worden de groepstests schriftelijk gerepresenteerd en worden de antwoorden door de onderzochten op de testformulieren zelf of op aparte antwoordformulieren ingevuld. Een andere mogelijkheid is dat de vragen via een beeldscherm worden gerepresenteerd. Ook een individuele test kan schriftelijk of via de computer worden voorgelegd, hiertegenover staat enerzijds de mondelinge test en anderzijds de verrichtingstest.

Type item

Indien items de vrije-antwoordvorm hebben moet de onderzochte zelf het antwoord bedenken en formuleren. Bij de keuze-antwoordvorm wordt de onderzochte verzocht een keuze te maken uit de vooraf door de testconstruër geformuleerde antwoordmogelijkheden. Andere termen zijn open versus gesloten vragen.