

Methoden van het wetenschappelijk onderzoek

(a) positief-wetenschappelijke benadering

door: Prof. Dr. K. Verfaillie

Hoofdstuk 1. De logica van het wetenschappelijk onderzoek

1.1. Wetenschap en gezond verstand

Er zijn twee criteria waaraan wetenschappelijke kennis moet voldoen

- a. **interne consistentie (coherentie)**
de uitspraak mag niet in tegenspraak zijn met de rest van de theorie waarvan de uitspraak deel uitmaakt
- b. **correspondentie met werkelijkheid (empirische verificatie)**
de verzamelde kennis wordt op haar overeenkomst met de ervaringswereld getoetst

Wetenschappelijke kennis wordt op een systematische en gecontroleerde manier verworven.

- a. **systematisch**: in overeenstemming met methodologische regels
- b. **gecontroleerd**: controle uitoefenen op storende variabelen (cf. infra)

Bij gebruik van gezond verstand worden theorieën niet op een systematische manier getoetst, er is sprake van selectiviteit. Mensen hebben de neiging evidentie te selecteren die consistent is met hun hypothese of met wat ze al weten.

Voorbeelden van selectiviteit:

- **heuristieken**: dit is een oplossingsstrategie die niet noodzakelijk tot de ‘juiste’ oplossing leidt. Men maakt hierbij vaak een mentale shortcut (vuistregel) om een conclusie te bereiken, die waarschijnlijk, maar niet noodzakelijk juist is. Meestal is deze manier nuttig, maar het kan tot fouten leiden.
- **beschikbaarheidsheuristiek**: wordt gebruikt om een oordeel te vellen over de waarschijnlijkheid van gebeurtenissen. Deze heuristiek is gebaseerd op de vraag “hoe gemakkelijk kan ik mij concrete voorbeelden voor de geest halen?”, hoe meer voorbeelden, des te meer beschikbaarheid. Maar de beschikbaarheid wordt ook door andere factoren dan frequentie en probabiliteit beïnvloed. Een klasse waarvan voorbeelden gemakkelijk uit het geheugen opgehaald kunnen worden zal groter *lijken* dan een klasse waarvan de voorbeelden moeilijker uit het geheugen op te halen zijn.
- **de rol van vividness van de informatie**: bepaalde informatie onthoudt je beter dan de daadwerkelijke statistische informatie. Veel mensen zullen bijvoorbeeld zeggen dat het gevaarlijker is met het vliegtuig te vliegen dan met de auto rijden, maar statistisch gezien is dit onjuist.
- **representativiteitsheuristiek**: dit komt vaak voor wanneer je moet schatten of geval X tot een bepaalde klasse behoort. Hoe goed lijkt dit geval op de andere gevallen in die klasse? Hoe representatief is dit geval voor deze klasse? Vaak is dit een bruikbare heuristiek, maar het houdt geen rekening met de basisprobabiliteit van een bepaalde klasse. Voorbeeld van de ingetogen Stef uit West-Vlaanderen, waarvan veel mensen dachten dat hij bibliothecaris was, maar eigenlijk was hij landbouwer. De basisprobabiliteit van landbouwers in West-Vlaanderen is groter dan die van bibliothecarissen.

1.2. Enkele opvattingen over de wetenschap

De **statische visie** op de wetenschap:

- nadruk op wetenschap als corpus van vaststaande feiten en theorieën
- aan verworven kennis wordt niet getornd; de taak van de wetenschapper bestaat erin deze kennis uit te breiden
- gevolg: wetenschappelijke veranderingen vinden plaats in de periferie
- werkt overmatig vertrouwen in de wetenschapper in de hand

De **dynamische visie** op de wetenschap:

- nieuwe kennis kan vroegere kennis wijzigen, zelfs radicaal wijzigen
- kan aanleiding geven tot compleet nieuwe theorieën, zelfs over dezelfde werkelijkheid
- wetenschappelijke veranderingen kunnen in centrum zelf plaatsvinden → wetenschappelijke revoluties (paradigmawissel)

Kanttekeningen bij deze visies:

- Enerzijds is de dynamische visie meer realistisch en legt ze meer nadruk op het belang van theorie, terwijl de statische visie de verzameling van feiten uit wil breiden. Anderzijds bestaat er de mythe dat wetenschap alleen bestaat uit dramatische ontdekkingen en doorbraken (= Einstein-syndroom).
- **Einstein-syndroom:**
 - het 'cruciale-experiment' model van wetenschappelijke vooruitgang; dit is een misvatting dat alle wetenschappelijke problemen opgelost worden in één enkel cruciaal experiment
 - het 'grote-sprong-voorwaarts' model van wetenschappelijke vooruitgang; dit is een misvatting dat theoretische vooruitgang het resultaat is van één enkel kritisch inzicht dat alle vorige kennis op de helling zet
 - cruciale experimenten en grote-theoretische-sprongen-voorwaarts komen wel voor, maar zijn niet de voornaamste, frequentste manier van vooruitgang maken
 - een meer accurate beschrijving van wetenschappelijke vooruitgang kan gemaakt worden met de begrippen *connectiviteitsprincipe* en *het principe van convergerende evidentie*.
- **Het connectiviteitsprincipe (Bronowski; cf. infra Lakatos):**
Om van vooruitgang te spreken moet een nieuwe theorie niet alleen nieuwe feiten verklaring, maar ook de oude
- **Het principe van convergerende evidentie:**
In plaats van te focussen op evidentie van één enkel zgn. cruciaal experiment, de evidentie van meerdere studies vergelijken en gebruiken, het liefst nog met verschillende onderzoeksmethodologieën

De **pragmatische visie** op de wetenschap:

- de wetenschap is gericht op de verbetering van onze leefwereld, op vooruitgang
- motief voor wetenschapper: oplossing zoeken voor een praktisch probleem

De **kennisgerichte visie** op de wetenschap:

- vooraleer we de leefwereld kunnen verbeteren, moeten we weten hoe de werkelijkheid van de leefwereld functioneert, welke wetmatigheden aan de basis liggen van dit functioneren
- motief voor de wetenschapper: een fenomeen begrijpen

Het onderscheid tussen de pragmatische en kennisgerichte visie komt overeen met het onderscheid tussen toepassingsgericht en fundamenteel onderzoek. De visies zijn echter wel complementair, kennisgericht en pragmatisch vullen elkaar aan.

De kennisgerichte visie legt de nadruk op verklaren. Dit verklaren geeft het vermogen om het verloop van gebeurtenissen te voorspellen. Dit kan heel nuttig zijn voor de toegepaste wetenschap. Het toegepaste onderzoek van de pragmatische visie kan relevante vragen aanbrengen die richtinggevend zijn voor vruchtbare ontwikkelingen in de fundamentele wetenschap.

Het zou echter fout zijn als alle wetenschap van in den beginne direct maatschappelijk relevant zou moeten zijn. De hoofdbedoeling van de wetenschap is niet het verbeteren van de mensheid, het is de bedoeling om kennis te verwerven en deze kennis vast te leggen in theorieën waarmee we fenomenen kunnen verklaren.

1.3. Doel van wetenschap: theorie

Een theorie is een verzameling **constructen (concepten)**, definities en stellingen die een systematische zienswijze geven op fenomenen, door **causale relaties** tussen **variabelen** te specificeren, met de bedoeling de fenomenen te **verklaren** en te **voorspellen** (Kerlinger, 1986).

Concept: Een theoretisch begrip dat een abstractie uitdrukt over de aanwezigheid van een eigenschap, dat verkregen is na een generalisatie over specifieke of particuliere observaties.

Construct: Een concept dat met opzet en bewust gecreëerd werd of aangepast werd voor wetenschappelijke doeleinden. Het verschil tussen concept en construct is dat een concept een abstractie over een eigenschap is door te generaliseren over specifieke objecten en een construct is een wetenschappelijk concept, zoals massa, energie, prestatiemotivatie en intelligentie.

Variabele: (Voorlopige definitie). Een variabele is een concept of construct dat verschillende waarden kan aannemen. Nog strikter genomen is een variabele een concept of construct dat verschillende numerieke waarden kan aannemen.

Causale relatie: Een oorzakelijk verband tussen X en Y. Een verandering in de variabele X leidt tot verandering in een andere variabele Y.

1. Theorie bestaat uit stellingen over gedefinieerde en gerelateerde constructen
2. Theorie specificiert de relaties tussen variabelen en geeft daardoor een systematische zienswijze op de fenomenen beschreven door de variabelen
3. Theorie verklaart fenomenen door te specificeren welke variabelen met welke variabelen gerelateerd zijn en hoe ze gerelateerd zijn. Dit stelt de onderzoeker in staat te voorspellen

De toetsing van de empirische verificatie van een theorie gebeurt in de empirische cyclus.

1.4. De empirische cyclus

1.4.1. De observatiefase

Het startpunt is een probleem, een obstakel of een idee. De onderzoeker ervaart een obstakel in het begrijpen van iets en is nieuwsgierig om te begrijpen. Meestal een observatie op een specifiek moment in een specifieke situatie die de interesse wekt. Toeval kan een rol spelen hierbij (*serendipisme/serendipiteit*). Dit betekent niet dat je alleen gelukt nodig hebt, het is ook nodig open te staan voor potentieel relevante observaties, de capaciteit om te zien dat een waarneming waardevol is. En de observatie wordt altijd al geleid door een theorie.

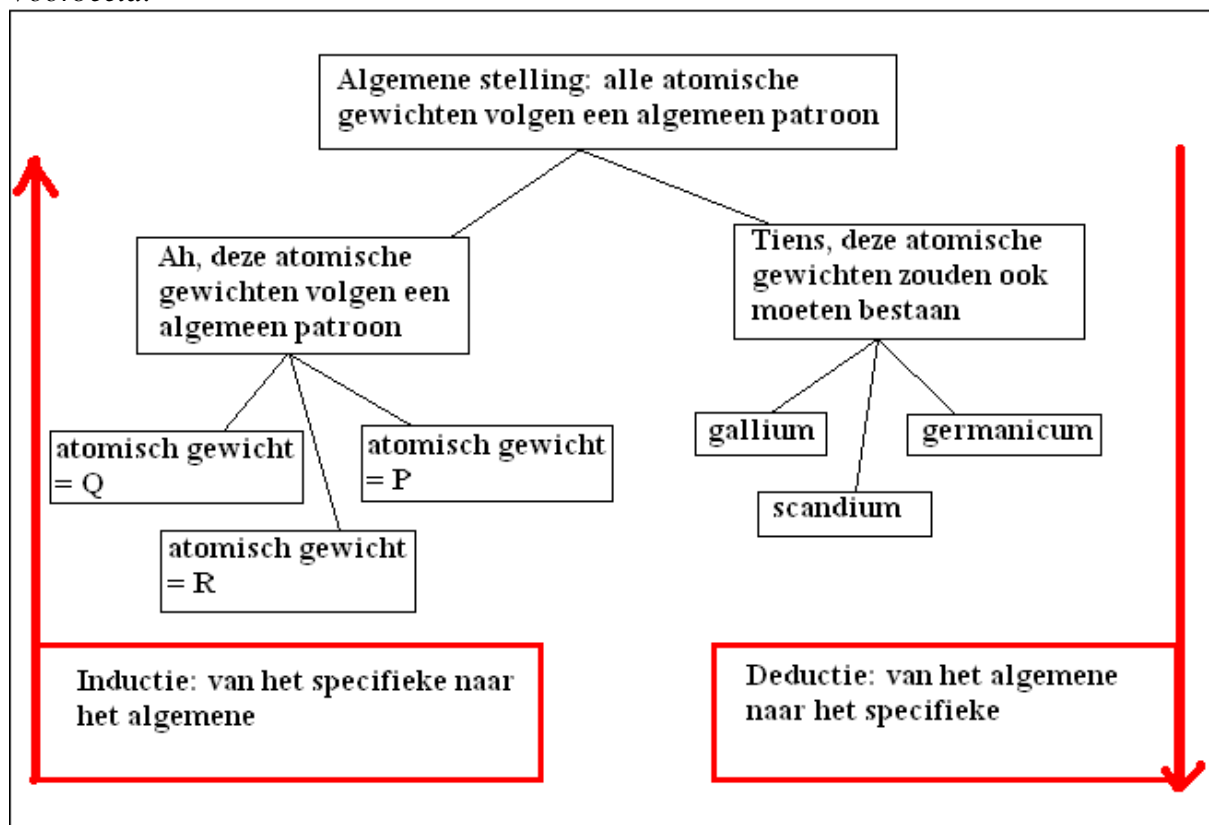
1.4.2. Inductieve fase

Het vergt soms heel wat denkwerk om het vaak vage gevoel uit de observatiefase om te zetten in een hanteerbaar probleem. Inductief houdt in dat vanuit particuliere observaties wordt gegeneraliseerd naar algemeen geldende uitspraken. Een stap die hierbij vaak over het hoofd wordt gezien is 'wat zijn relevante problemen in een wetenschap?'. Daarbij komt dat wanneer een probleem goed gedefinieerd is je vaak al dichterbij de oplossing bent.

1.4.3. Deductieve fase

Vanuit een theorie worden voorspellingen gemaakt, een hypothese en/of verwachting wordt opgesteld. Deductief houdt in dat het zich afspeelt binnen het kader van de theorie zelf (op abstract redeneringsniveau). Dit mondt uit in predicties, die voor empirische toetsing vatbaar zijn.

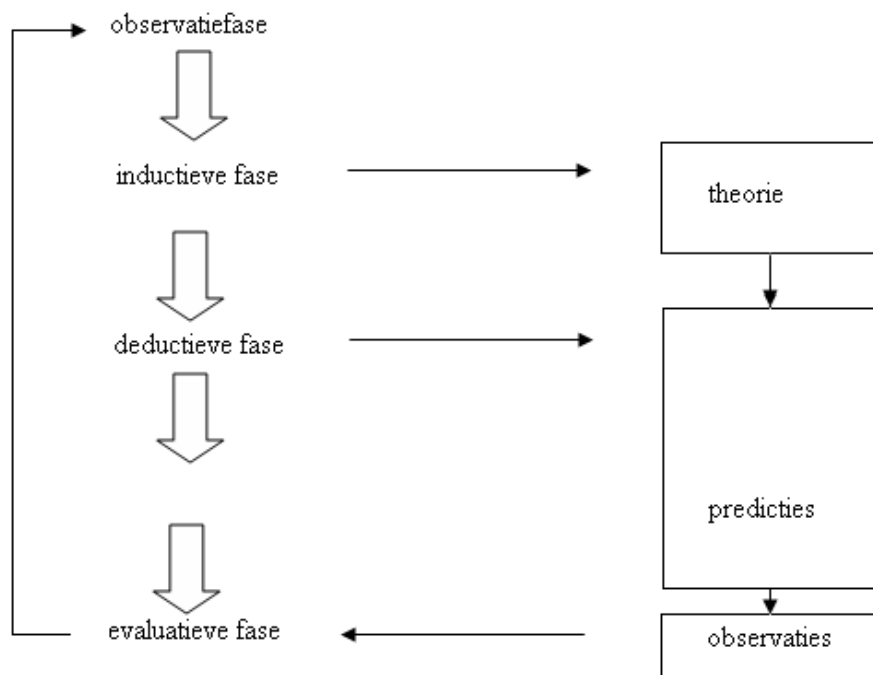
Voorbeeld:



1.4.4. Evaluatieve fase

Wanneer theorie X juist is zal hypothese H bevestigd worden. Als de hypothese H verworpen wordt, dan is de theorie X verkeerd en zal deze moeten worden aangepast. Wanneer hypothese H wordt bevestigd, dan zou theorie X juist kunnen zijn. Maar let op, dit is nog geen absolute zekerheid.

De empirische cyclus



1.5. Operationaliseren van variabelen

Positieve bekrachtiger Gebeurtenis die de kans op respons verhoogt wanneer ze op de respons volgt. Bijvoorbeeld wanneer je een rat voedsel geeft nadat ze op een pedaal heeft gedruwd, dan is voedsel een positieve bekrachtiger voor op het pedaal duwen.

Negatieve bekrachtiger Gebeurtenis die de kans op respons verhoogt wanneer ze verdwijnt na de respons. Bijvoorbeeld een moeder die reageert op het huilen van een baby. De rust die daardoor ontstaat is een negatieve bekrachtiger voor de moeder om het gedrag te herhalen.

De sterkte waarmee een respons geassocieerd is met een stimulus neemt toe, naarmate het aantal bekrachtigingen onmiddellijk na de respons toeneemt en naarmate de intensiteit van elke bekrachtiging toeneemt. De twee variabelen werken onafhankelijk van elkaar.

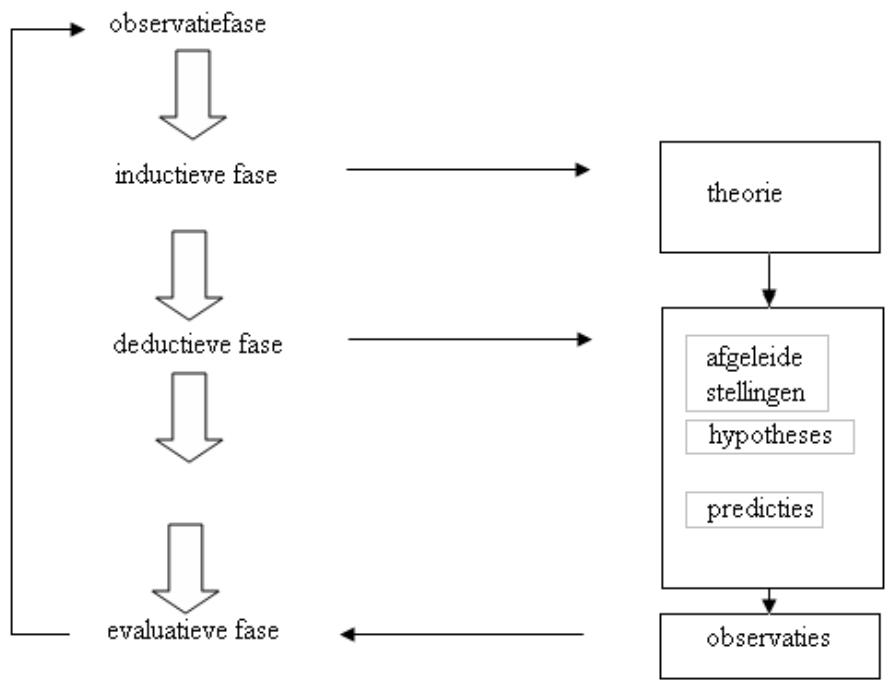
Stellingen zijn beweringen over hoe de werkelijkheid functioneert, door causale relaties tussen concepten te specificeren. Ze bezitten een universele geldigheid en ze zeggen nog niet hoe de onderzoeker dit concreet in de werkelijkheid zou kunnen nagaan.

Een afgeleide stelling (*deductie*) kan door logisch redeneren uit een theorie afgeleid worden. Maar ook deze zegt niet hoe dit concreet in de werkelijkheid nagegaan moet worden. Bij een hypothesestelling wordt de stelling geherformuleerd in termen van observeerbare (en manipuleerbare) variabelen. De hypothese verwijst naar gebeurtenissen die in principe observeerbaar zijn, maar ze is nog steeds geformuleerd als een uitspraak met een universele geldigheid. Vervolgens bedenk je een concrete situatie die voldoet aan de voorwaarden gespecificeerd in de hypothese en dan ga je na of het veronderstelde verband inderdaad geobserveerd wordt. Dit gebeurt in de vertaling van hypothese naar predictie.

Het onderscheid tussen afhankelijke en onafhankelijke variabele

Een onafhankelijke variabele is een variabele die gemanipuleerd wordt. De onafhankelijke variabele is diegene waarvan verwacht wordt dat ze een invloed heeft op het gedrag. Het aantal fouten dat iemand maakt bij hun zesde poging door het doolhof is een afhankelijke variabele. De afhankelijke variabele is de variabele die geobserveerd wordt en wordt beïnvloed door of is afhankelijk van de onafhankelijke variabele.

Als laatste stap voer je het experiment concreet uit om na te gaan of de observaties overeenkomen met de predicties. Als je predicties worden bevestigd, dan is de theorie (voorlopig) houdbaar. Worden ze niet bevestigd, dan moet de theorie herzien worden.



1.6. Hypotheses

- a. Er bestaat een verband tussen het aantal afleiders en de reactietijd: wanneer het aantal afleiders wijzigt, zal ook de reactietijd wijzigen.
- b. Er bestaat een stijgend verband tussen het aantal afleiders en de reactietijd: wanneer het aantal afleiders toeneemt, zal ook de reactietijd toenemen.
- c. Er bestaat een stijgend lineair verband tussen het aantal afleiders en de reactietijd: bij een even grote toename van het aantal afleiders, is er een even grote toename in reactietijd.

Een bruikbare hypothese voldoet aan twee voorwaarden:

1. Er wordt expliciet verwezen naar het bestaan van een verband tussen variabelen.
2. De hypothese verwijst al naar de mogelijkheid van empirische toetsing.

We onderscheiden verschillende soorten hypothesen:

- De hypothese in de nulvorm: een hypothese die stelt dat er geen verband is tussen variabelen.
- De hypothese die stelt dat er een verband is, zonder zich uit te spreken over de richting van het verband.
- De hypothese die zich ook uitspreekt over de richting van het verband.
- De hypothese die zich ook uitspreekt over de specifieke vorm en grootte van het verband.

NB. Normaliter is een onderzoekshypothese geen nulhypothese, een theorie kan er namelijk weinig baat bij hebben te specificeren waar er allemaal geen verband zou bestaan.

Bij onderzoek kun je ook de aandacht toewijzen; parallelle vs. seriële verwerking.

Je kunt de proefpersoon laten zoeken naar een enkelvoudig kenmerk, zoals bijvoorbeeld een rood streepje. De doelstimulus kan bij een enkelvoudig kenmerk van de afleiders onderscheiden worden op basis van één enkel kenmerk. Hier vindt dan pre-attentieve, parallelle verwerking plaats, want het lijkt alsof de waarnemer alle locaties tegelijk/parallel kan onderzoeken. Je verkrijgt hierbij een vlakke zoekcurve.

Je kunt de proefpersoon ook laten zoeken naar een conjunctie van kenmerken, zoals bijvoorbeeld een rood, naar rechts georiënteerd streepje. De doelstimulus kan alleen van afleiders worden onderscheiden door een combinatie van kenmerken. Hier vindt dan attentieve, seriële verwerking plaats, want het lijkt alsof de waarnemer elke locatie één voor één moet onderzoeken. Je verkrijgt hierbij een zoekcurve met een stijgend lineair verband.

Dit soort onderzoeken laat toe de 'primitieve' kenmerken van perceptie te ontdekken. Wanneer je een stimulus kunt vinden onafhankelijk van het aantal afleiders, dan is dat vermoedelijk een primitief kenmerk, zoals bijvoorbeeld de kromming van een lijnstuk, de oriëntatie van een lijnstuk, de kleur en helderheid en de beweging.

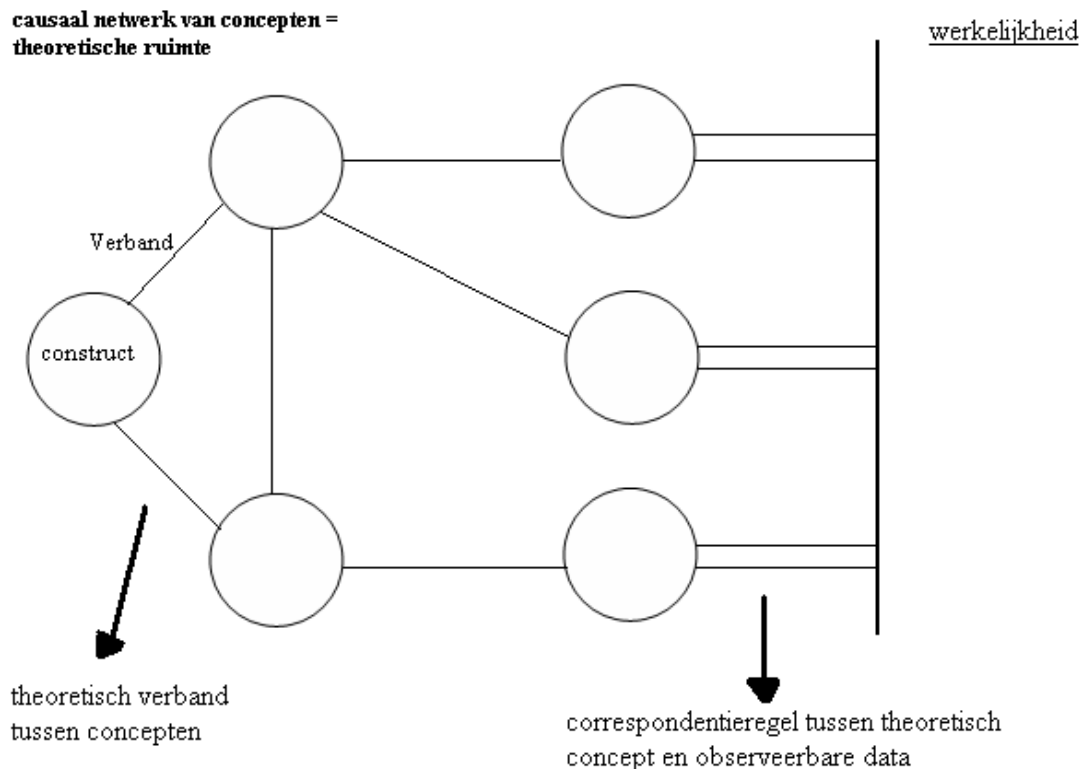
Soms kan een nulhypothese dus wel zinvol zijn, omdat we in het geval van 'primitieve' kenmerken geen verband voorspellen tussen het aantal afleiders en de reactietijd.

1.7. Relatie theorie-empirische cyclus

1.7.1. Inleiding

Theorie: Een causaal netwerk van theoretische constructen.
Correspondentiecriterium: Delen van de theorie moeten overeenkomen met de werkelijkheid.

Wetenschap bestaat uit theorie aan de ene kant en empirische evidentie aan de andere kant.



1.7.2. Onderscheid model-theorie

Model:

- Causaal netwerk van theoretische concepten
- Netwerk van verbanden tussen theoretische concepten staat "model" voor datgene wat er in de werkelijkheid gebeurt.
- Model kan getoetst worden op logische gronden (cf. coherentiecriterium)
- Nog niet met de werkelijkheid verbonden, maar staat wel als model voor de werkelijkheid

Coherentiecriterium: rol van computersimulatie. Het nabootsen van dit geheel van deelprocessen in een computerprogramma. Een voordeel is dat de onderzoeker elk deel proces moet expliciteren. Bij inconsistenties of onvolledigheden zal het programma niet werken.

Theorie:

- Model waarbij bepaalde constructen verbonden zijn met de empirische wereld door correspondentieregels
- Correspondentieregel geeft de procedure aan om een construct aan

- de data te relateren
- De theorie kan empirisch getoetst worden

1.7.3. Twee verschillende definities van constructen

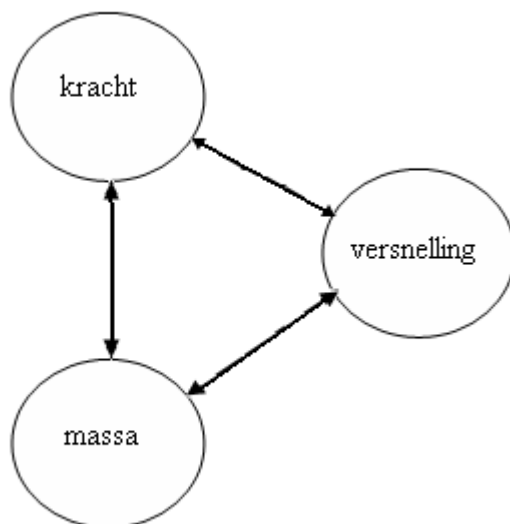
Operationele definitie

Het construct wordt gedefinieerd in termen van observeerbare data en/of manipuleerbare variabelen. Deze komt overeen met de dubbele lijnen in de figuur. Het is dus de verbinding met de werkelijkheid.

- Voorbeelden:
- Gewicht wordt gedefinieerd in termen van concrete wegingsoperaties
 - Tijd definiëren door te refereren naar klokken
 - Intelligentie wordt bepaald door aan te geven welke gedragingen we intelligent vinden

Constitutieve definitie

Het construct wordt gedefinieerd in termen van andere constructen in de set, min of meer zoals in een woordenboek gebeurt. In de figuur is dit dus het netwerk van enkelvoudige lijnen. Deze is per definitie circulair.



Constitutieve definitie

$kracht = massa \times versnelling$
 $massa = kracht / versnelling$
 $versnelling = kracht / massa$

Een construct kan verschillende operationele definities hebben. Welke instructies moet ik volgen om iets te observeren?

Een construct kan ook verschillende constitutieve definities hebben. Wat is de definitie van het construct in termen van andere constructen? Dit is afhankelijk van de concepten waarmee het in verband wordt gebracht.

1.7.4. Latente en manifeste variabelen

Latente variabele: Deze variabele is onzichtbaar en wordt ook wel de constructvariabele (constitutief gedefinieerd) genoemd.

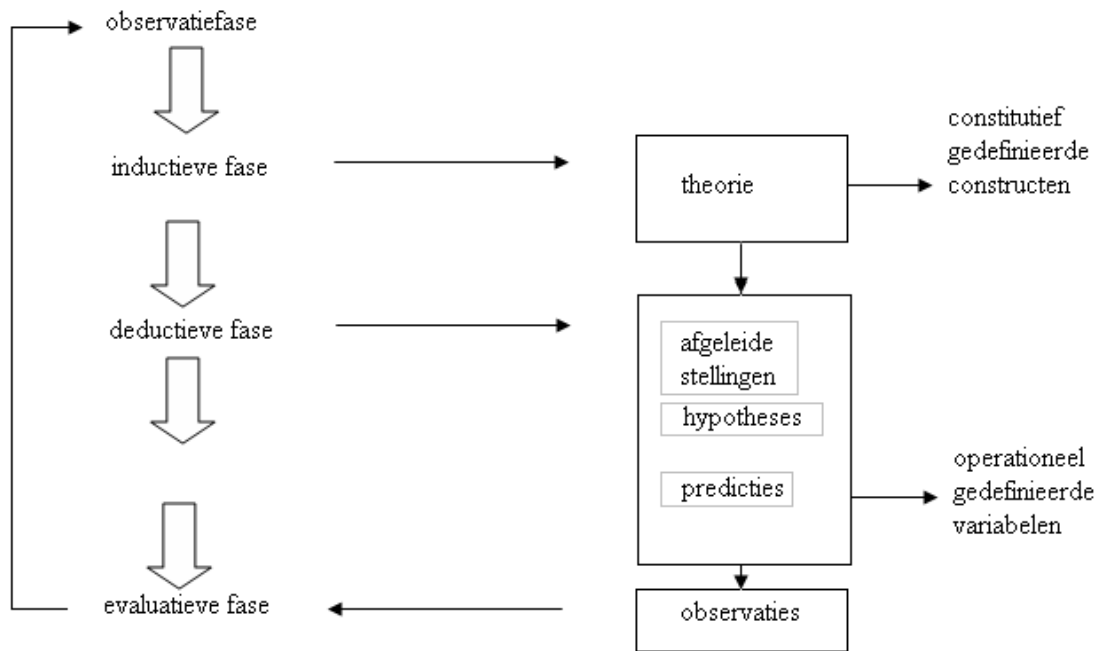
Vb.: Mensen verschillen in intelligentie. Mensen situeren zich verschillend op een variabele die niet rechtstreeks zichtbaar is. De variabiliteit in de latente variabele is per definitie niet direct

observeerbaar.

Manifeste variabele: Deze variabele is de observeerbare variabiliteit en wordt ook wel indicatorvariabele of observatievariabele (operationeel gedefinieerd) genoemd.

Vb.: Mensen verschillen wanneer zij taken moeten oplossen waarvan men zegt dat men over een zekere mate van intelligentie moet beschikken om ze te kunnen oplossen.

1.7.5. Verband met operationalisering



1.7.6. Extreem operationalisme

Volgens het extreem operationalisme moeten alle constructen een operationele definitie hebben. Het extreem operationalisme is fout. Alle constructen moeten wel tenminste indirect (via hun connectie met andere constructen) met observeerbare data verbonden zijn. Wanneer dit niet zo is, heeft het construct geen verklarende kracht.

Een concept wordt het best operationeel gedefinieerd door een set van operaties (= notie van convergerende operaties). Uit de modelmatig veronderstelde verbanden tussen concepten (zelfs als ze niet allemaal operationeel te definiëren zijn) kunnen predicties afgeleid worden die voor empirische toetsing vatbaar zijn.

1.7.7. Correlationele vs. theoretische verklaringen

Correlationele procedures of verklaringen behoren tot een wetenschap die voornamelijk bestaat uit stellingen die de relatie weergeven tussen variabelen die min of meer direct observeerbaar zijn.

Theoretische procedures of verklaringen behoren tot een wetenschap die ook relaties verklaart op basis van principes die niet onmiddellijk gegeven zijn en verder gaan dan strikte empirische kennis.

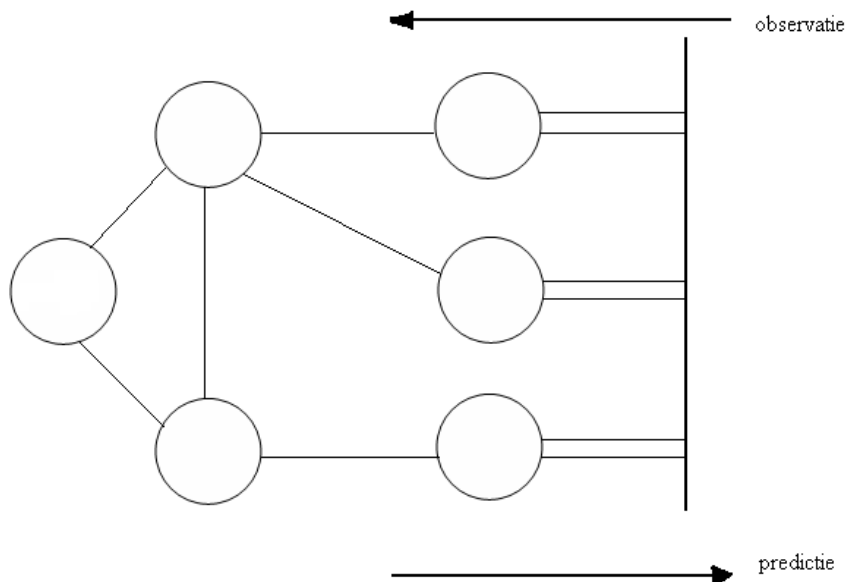
Let op!: Geen enkele wetenschap is puur correlatief of puur theoretisch. Er is wel een verschil in de mate waarin ze correlatief of theoretisch zijn. Elke wetenschap begint als puur correlatief en wordt naarmate de tijd vordert meer en meer theoretisch.

Fysica is een voorbeeld van een “sterke” wetenschap. Er zijn veel constructen, veel correspondentieregels en veel relaties tussen constructen. De correspondentieregels zijn vaak operationele definities van hoeveelheden waardoor aan de constructen getallen toegekend worden. De theoretische constructen zijn vaak ook kwantificeerbaar. Ook kan de relatie tussen de constructen de vorm van een mathematische vergelijking aannemen.

Psychologie is een voorbeeld van een minder “sterke” wetenschap. Enerzijds zijn de theoretische constructen belangrijk voor het verklaren van fenomenen in de psychologie. De connecties zijn vaak vaag en verbaal uitgedrukt i.p.v. mathematisch, zoals bij de fysica. Anderzijds heb je de operationeel gedefinieerde constructen. Het probleem is dat men het er vaak niet over eens is dat de operationeel gedefinieerde concepten hetzelfde zijn als de theoretisch gedefinieerde concepten. Het gevolg voor de empirische cyclus is dat wanneer de predictie niet bevestigd wordt de theorie of fout wordt verklaard of dat de operationeel gedefinieerde variabelen geen goede metingen zijn van de corresponderende constitutief gedefinieerde variabelen.

1.7.10. Verband met de empirische cyclus

De correspondentieregels werken in twee richtingen. De observatie van empirie naar construct en de predictie van construct naar empirie.



Wanneer relaties tussen constructen mathematische vergelijkingen zijn, noemen we dit formele modellen (zoals in de fysica). Een formeel kan data a.h.w. berekenen in plaats van de echte observaties te doen. Het kan ook een procedure zijn om een theorie te testen: worden de predicties door de observaties bevestigd of niet?

1.9. Beoordelingscriteria voor theorieën

- Een uitspraak mag niet in tegenspraak zijn met de rest van de theorie waarvan de uitspraak deel uitmaakt.
- Tenminste een deel van de theorie moet overeenkomen met de ervaringswereld

Deze twee voorwaarden zijn de noodzakelijke beoordelingscriteria voor een theorie.

Extensiviteit

Hoe wijdomvattend is de theorie? Hoe omvangrijk is het specifieke domein binnen de observeerbare werkelijkheid die een theorie beweert te kunnen verklaren?

Spaarzaamheid

Bij de keuze tussen twee theorieën die even veel verklaren (met dezelfde extensiviteit), maar die verschillen in complexiteit, kies dan voor de meest eenvoudige theorie. Het spaarzaamheids principe wordt ook wel Ockham's razor genoemd.

Extensiviteit en spaarzaamheid kunnen met elkaar in conflict komen. Enerzijds is de theorie sterker wanneer de empirische inhoud kan uitgebreid worden zonder het verklarende begrippenkader uit te breiden. Anderzijds is het vaak naar aanleiding van een poging om theorie te laten winnen aan extensiviteit dat het begrippenkader dient uitgebreid te worden.

Maatschappelijke relevantie

Een theorie wordt interessanter naarmate ze toelaat er meer praktische toepassingen uit af te leiden. Zeker in het geval van een bijdrage tot het oplossen van sociaal-economische, medische of technische problemen.

1.10. De evaluatieve fase

Karl Popper (°1902 - +1994)

Het empirisch onderzoek legt te sterk de nadruk op het bevestigen van de gestelde hypothese door de waarnemingen.

Een reden om achterdochtig te zijn bij zo'n strategie zijn dat het vooringenomenheid (1) in de hand werkt. Mensen zoeken informatie die consistent is met hun hypothese en ze vermijden potentieel falsifiërende informatie. Door de keuze van de proefopzet gebeurt er al een selectie in wat men kan observeren en wat niet. Het observeren zelf kan ook nog eens selectief zijn.

Dan is er nog de invloed van kennis waarover men al beschikt die de perceptie beïnvloed. Het verschil tussen bottom-up en top-down verwerking is er ook nog. Bottom-up is datagedreven, top-down is conceptueel gedreven. Op basis van vroeger opgedane kennis gaat het visuele systeem hypothesen genereren over voorwerpen die aanwezig kunnen zijn in de wereld en gaat vervolgens op zoek naar evidentie om deze hypothese te bevestigen/verwerpen.

Een tweede reden is het logisch bezwaar (2). Een eenmalige bevestiging van een predictie aan de hand van geobserveerde gegevens biedt geen absolute zekerheid over de algemene

geldigheid van de hypothese waaruit de predictie werd afgeleid. Zelfs bij herhaaldelijke waarnemingen is er nooit volledige zekerheid.

Popper haal hieruit de conclusie dat het beter is te werken met theorieën die tot falsifieerbare hypothesen leiden (Popperiaans falsificationisme).

Van een theorie moet tenminste aangetoond kunnen worden hoe ze verworpen zou kunnen worden. Toetsbare theorieën zijn interessanter, want een theorie zegt meer in de mate dat ze zich openstelt voor weerlegging en is des te waardevoller als de theorie tegen het risico van weerlegging bestand is.

Er zijn echter wel wat problemen wanneer je het Popperiaans falsificationisme consequent doortrekt. De theorieën zullen zich namelijk bescheidener opstellen. Het kan ook niet de enige bedoeling van empirische verificatie om theorieën te proberen te verwerpen.

Lakatos (1970)

Lakatos was van het genuanceerd falsificationisme. Hij zei dat een theorie (T) verlaten kan worden ten voordelen van een aangepaste versie van die theorie (T') indien:

1. T' niet de gefalsifieerde empirische inhoud van T omvat. T' verklaart dus tenminste wat T al verklaarde.
2. T' heeft een grotere empirische inhoud dan T. T' verklaart dus meer dan T en het is zelfs mogelijk dat T' fenomenen voorspelt die volgens T onwaarschijnlijk of zelfs onmogelijk zijn.
3. Een gedeelte van de meerinhoud van T' wordt niet door de werkelijkheid gefalsificeerd.

Opmerking: T' zal op den duur moeten wijken voor T'' → dynamische visie op wetenschap

Objectiviteit door consensus

Hoe beslissen wetenschappers of observaties al dan niet overeenstemmen met de werkelijkheid? Dit doen ze door selectiviteit en top-down invloeden op de waarneming. Ook door objectiviteit door consensus: observaties worden in één richting geïnterpreteerd indien iedereen het over de interpretatie eens is.

De objectiviteit door consensus is echter niet waterdicht. Er moet ook na worden gedacht over het feit of de theorie wel realistisch is. Het is daarom belangrijk om regels te ontwikkelen over wat ondersteunende en falsifiërende evidentie is en hoe we consensus bereiken over welke evidentie.

Hoofdstuk 2. Meten in de psychologie

2.1. Inleiding

Wetenschappelijke kennis wordt in eerste instantie opgebouwd door uitspraken te doen over de onderlinge samenhang tussen variabelen. In de psychologie ondervinden we hierbij twee problemen.

1. Het probleem van de meetbaarheid van psychologische variabelen. In welke mate heeft het zin om te spreken over fundamenteel meten in de psychologie?
2. Hoe kunnen we causaliteit uit geobserveerde gegevens infereren? Hoe kunnen we uit gegevens valide conclusies trekken over de onderzoekshypothese? Hoe kunnen we de zekerheid verhogen dat er geen storende variabele in het spel is?

2.2. Meetniveaus

Natuurlijke variabelen

Een natuurlijke variabele is een indeling van alle mogelijke onderzoeksobjecten op basis van een welbepaalde eigenschap in een aantal wederzijds uitsluitende klassen, waarbij ieder object tot één en slechts één klasse behoort. De klassen komen overeen met de waarden die de variabele aan kan nemen.

- variabelen: de personen indelen op basis van geslacht, aardbevingen indelen o.b.v. hun hevigheid, voorwerpen indelen o.b.v. hun temperatuur, etc.
- onderzoeksobjecten: personen, aardbevingen, voorwerpen, etc.
- eigenschap: geslacht, hevigheid, temperatuur, etc.
- waarden: mannelijk of vrouwelijk, ... op de schaal van Richter, ... graden Celsius, etc.

Het studieobject van de psychologie is het gedrag en de onderliggende mentale processen. De psychologie bestudeert de variabiliteit in gedrag en mentale processen in functie van verschillen in situaties, personen (en tijd). Je geeft aan welke aspecten van de situatie en gedrag kunnen veranderen en die aspecten komen in aanmerking om als variabele te fungeren.

Discrete vs. continue variabele

Bij een discrete (categorische) variabele is elk punt op de schaal volledig gescheiden van het volgende. Om van de ene waarde naar de andere te gaan moet een sprong gemaakt worden, de waarde kan dus slechts stapsgewijs veranderen en in die zin is de discrete variabele equivalent met natuurlijke getallen. Discrete variabelen kunnen verschillen in de fijnheid waarmee een klassenindeling wordt gemaakt. De meeteenheid voor het meten van de hoogte van de Eiffeltoren is een minder fijne klassenindeling dan wanneer je de lengte van een wijsvinger meet.

Bij een continue variabele is er geen grens op de onderverdeling van punten. Tussen elke twee waarden kan een grens gevonden worden en in die zin is ze dus equivalent met reële getallen. De continuïteit is een theoretische veronderstelling en omdat het meetinstrument altijd beperkt is kunnen eigenlijk alleen discrete variabelen geobserveerd worden. In praktijk spreken we toch van een continue variabele wanneer de variabele een groot aantal waarden kan aannemen en de variabele geconceptualiseerd kan worden als manifestatie van een onderliggende continue variabelen. Hoe fijner de klassenindeling van een discrete variabele, hoe dichter we in de buurt komen van de overgang van discrete naar continue variabele.

Van waarden naar meetwaarden

Indien men de waarden van de variabele voorziet van getalwaarden, krijg je een geschaalde variabele of een schaal. Deze getallen zijn meetwaarden. Meten is het toekennen van getallen aan de waarden van een natuurlijke variabele, zodanig dat de geobserveerde relaties tussen de waarde afgebeeld zijn in overeenkomstige relaties tussen de getallen. De getallen geven eigenlijk dezelfde informatie als de klasse-indeling en zijn de meetwaarden van de objecten m.b.t. de variabele.

Een *relationeel systeem* is een verzameling objecten en één of meerdere relaties gedefinieerd tussen de objecten. In een *empirisch relationeel systeem* zijn de objecten en relaties empirisch, zoals bijvoorbeeld mensen en gewicht. Objecten zijn de entiteiten waarop de meting betrekking heeft en relaties zijn de observeerbare relaties tussen entiteiten, zoals bijvoorbeeld de grootte. In een *numerisch relationeel systeem* zijn de objecten getallen.

Meten is het afbeelden van een empirische relationeel systeem in een numerisch relationeel systeem. Het veronderstelt een homomorfisme tussen de twee systemen. Bij meten bestaat er een vrijheid in het toekennen van de getallen, de enige voorwaarde is dat de relaties tussen elementen in het empirisch relationeel systeem worden weerspiegeld in relaties in het numerisch relationeel systeem. Dit noemen we transformatievrijheid. Na de transformatie moeten de nieuwe getallen wel dezelfde informatie geven als de oorspronkelijke getallen.

Verschillende meetniveaus hebben eigenlijk te maken met de relaties van het empirisch relationeel systeem die weerspiegeld moeten worden in het numerisch relationeel systeem.

Nominaal niveau

Elk item/voorwerp/gebeurtenis moeten we in één enkele categorie kunnen plaatsen. Dit zijn mutueel exclusieve en exhaustieve categorieën. De leden van eenzelfde categorie krijgen hetzelfde label. Deze labels plakken noemt men het nominaal niveau, omdat de labels slechts een naam zijn, maar dit betekent niet dat het niet numerisch kan zijn. De getallen die op de categorieën geplakt worden gebruiken we echter niet als echte rekenkundige getallen.

De relatie tussen objecten in het empirisch relationeel systeem: behoort een onderzoeksobject voor een bepaalde eigenschap tot dezelfde klasse of niet? De relatie tussen objecten in het numerisch relationeel systeem: getallen kunnen verschillend zijn. Dus: je kent de getallen toe aan de klassen op zo'n manier, dat onderzoeksobjecten die tot eenzelfde klasse behoren ook hetzelfde getal toebedeeld krijgen en objecten die tot een verschillende klasse behoren dus een verschillend getal.

Ordinaal niveau

Op ordinaal niveau wordt er informatie over de rangorde gegeven. Ordinale getallen stellen de volgorde voor van de onderzoekselementen. Een voorbeeld hiervan is de schaal van Richter en de schaal van Beaufort.

De relatie tussen objecten in het empirisch relationeel systeem: onderzoeksobjecten kunnen geordend worden naargelang hun waarde voor een bepaalde eigenschap. De relatie tussen objecten in het numerisch relationeel systeem: van twee verschillende getallen is het ene getal groter dan het andere. Dus je kent de getallen zo toe aan de klassen dat de onderzoeksobjecten die een hogere rang hebben ook een hoger getal toebedeeld krijgen.

Geordend metrisch niveau

De relatie tussen objecten in het empirisch relationeel systeem: onderzoeksobjecten kunnen geordend worden naargelang hun waarde voor een bepaalde eigenschap + paren van onderzoeksobjecten kunnen geordend worden naar onderlinge gelijkheid op de eigenschap. De relatie tussen objecten in het numerisch relationeel systeem: van twee verschillende getallen is het ene getal groter dan het andere ($A > B$) + van twee verschillende verschillen tussen getallen is het ene verschil groter dan het andere. Dus je kent de getallen zo toe aan de klassen dat de onderzoeksobjecten die een hogere rang hebben ook een hoger getal toebedeeld krijgen en dat grotere verschillen tussen onderzoeksobjecten overeenkomen met grotere verschillen tussen getallen.

Niet elke monotoon stijgende transformatie zal alle oorspronkelijke informatie behouden. Alleen een hypermonotoon stijgende transformatie zal alle informatie behouden.

Intervalniveau

Het heeft op intervalniveau zin om over de verhouding van afstanden te spreken. Op intervalniveau kunnen we op een zinvolle manier over afstanden tussen punten op een schaal spreken. De afstanden (of intervallen) tussen onderzoekselementen (m.b.t. een bepaald kenmerk) worden weerspiegeld in de afstanden tussen getallen.

De relatie tussen objecten in het empirisch relationeel systeem: de onderzoeksobjecten kunnen geordend worden naargelang hun waarde voor een bepaalde eigenschap + relatieve afstanden (verhoudingen van telkens twee verschillen tussen onderzoeksobjecten op die eigenschap) hebben betekenis. De relatie tussen objecten in het numerisch relationeel systeem: van twee verschillende getallen is het ene getal groter dan het andere ($A > B$) + getalsverhoudingen van twee verschillen hebben een bepaalde grootte. Dus je kent de getallen zo toe aan de klassen dat de onderzoeksobjecten die een hogere rang hebben ook een hoger getal toebedeeld krijgen en dat relatieve afstanden (of intervallen) tussen onderzoekselementen (m.b.t.) een bepaald kenmerk worden weerspiegeld in de afstanden tussen getallen.

Transformatievrijheid op intervalniveau: De transformaties moeten de relatieve afstanden niet beïnvloeden, zoals bijvoorbeeld de transformatie van de Celsiusschaal naar Fahrenheit-schaal.

Rationiveau of verhoudingsniveau

We hebben een echt nulpunt (een natuurlijke ondergrens) voor de waarden van een variabele. De negatieve getallen hebben geen betekenis. De meeteenheid is arbitrair. Het gevolg hiervan is dat het zin heeft om over de verhoudingen tussen onderzoekselementen (m.b.t. een bepaald kenmerk) te spreken.

De relatie tussen objecten in het empirisch relationeel systeem: onderzoeksobjecten kunnen geordend worden naar gelang hun waarde voor een bepaalde eigenschap + verhoudingen van onderzoeksobjecten op die eigenschap hebben een betekenis. De relatie tussen objecten in het numerisch relationeel systeem: van twee verschillende getallen is het ene getal groter dan het andere ($A > B$) + getalsverhoudingen hebben een bepaalde grootte. Dus je kent de getallen zo toe aan de klassen dat de onderzoeksobjecten die een hogere rang hebben ook een hoger getal toebedeeld krijgen en dat de verhoudingen tussen onderzoekselementen (m.b.t. een bepaald kenmerk) worden weerspiegeld in de verhoudingen tussen de getallen.

Transformatie op rationiveau: De transformaties moeten de verhouding van elk tweetal van schaalwaarden niet beïnvloeden, zoals bijvoorbeeld de transformatie van guldens naar euro's.

Absoluut niveau

De getallen voor de waarden van een variabele liggen geheel vast, dit betekent dat ook de oorsprong en de meeteenheid vastliggen.

Transformatievrijheid op absoluut niveau: je hebt eigenlijk alleen de identiteitstransformatie, waarbij je het dus over een totaal andere variabele gaat hebben. Op absoluut niveau is er eigenlijk maar één schaal voor de variabele.

Hiërarchie van meetniveaus

Schalen

- absolute schalen zijn ook verhoudingsschalen, intervalschalen, geordend metrische schalen, ordinale schalen en nominale schalen
- verhoudingsschalen zijn ook intervalschalen, geordend metrische schalen, ordinale schalen en nominale schalen
- intervalschalen zijn ook geordend metrische schalen, ordinale schalen en nominale schalen
- geordend metrische schalen zijn ook ordinale schalen en nominale schalen
- ordinale schalen zijn ook nominale schalen

Transformatievrijheid

- meten op een bepaald niveau legt voorwaarden op aan de geobserveerde gegevens, waarop men zich baseert om een variabele te meten
- die voorwaarden worden strikter naarmate we van nominaal naar rationiveau gaan
- transformatievrijheid wordt kleiner:

Absoluut niveau	→	Identiteitstransformatie
Verhoudingsniveau	→	Similariteitstransformatie
Intervalniveau	→	Positief lineaire transformatie
Geordend metrisch niveau	→	Hypermonotoon stijgende transformatie
Ordinaal niveau	→	Strikt monotoon stijgende transformatie
Nominaal niveau	→	Eén-één transformatie

Een transformatie is telkens een deelverzameling van de onderliggende transformatie.

Informatie

Naarmate we van nominaal naar absoluut gaan krijgen we telkens meer specifieke informatie.

Nominaal	→	informatie over (on)gelijkheden
Ordinaal	→	ook informatie over rangordening
Geordend metrisch	→	ook informatie over rangordening van onderlinge afstanden
Interval	→	ook informatie over verhouding van twee verschillen
Ratio	→	ook informatie over verhouding van twee schaalwaarden

Verband met de theorie

Eisen over het meetniveau gaan terug op het soort inhoudelijke uitspraken die uit een theorie kunnen worden gededuceerd. Het heeft ook praktische consequenties, de specificaties van een hypothese zorgen dat het nuttig is voor het opstellen van een werkplan.

2.3. Inleiding in datatheorie

Ook in de datatheorie beeldt men het empirisch systeem af op het numerisch systeem. Daarbij is het vaak handig om beroep te doen op de isomorfie tussen reële getallen en punten op een rechte lijn. Dit deden we soms al bij de bespreking van de meettheorie, maar hier beelden we de elementen van het empirisch systeem nog explicieter af op getallen in het numerisch systeem (numerische modelformulering). Dit is gelijk aan het identificeren van elementen uit het empirisch systeem als punten op een rechte lijn (geometrische modelformulering). Bij een geometrische modelformulering wordt elk element voorgesteld als één punt op een lijn in een twee- of meerdimensionale ruimte.

Coombs onderscheidt vier fundamenteel verschillende types van data. Enkelvoudige prikkelgegevens, voorkeurgegevens, prikkelvergelijkingsgegevens en vergelijkingen van verschillen tussen prikkels.

Enkelvoudige prikkelgegevens

Er zijn twee formele kenmerken van enkelvoudige prikkelgegevens:

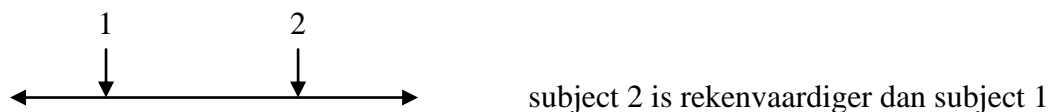
- twee verzamelingen van entiteiten worden als twee afzonderlijke verzamelingen van punten in de geometrische schaaloplossing voorgesteld.
- de geobserveerde relatie tussen twee punten (een entiteit uit de ene verzameling en een entiteit uit de andere verzameling) wordt voorgesteld in een relatie tussen twee punten.

Voorbeeld met een rekenvaardigheidstest:

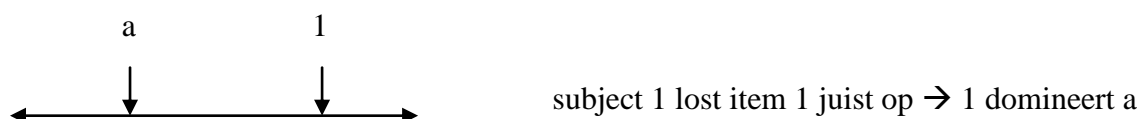
Elk item heeft een bepaalde moeilijkheidsgraad. Je stelt het moeilijkheidscontinuüm voor als een rechte lijn. Elk item kan als een punt op de lijn gesitueerd worden naargelang zijn moeilijkheidsgraad.



Je kunt ook twee subjecten vergelijking m.b.t. hun rekenvaardigheid.



De relatie subject-item kan ook weergegeven worden. Deze relatie is een dominantierelatie, dit houdt in dat wanneer het subject het item juist oplost zij dominant is aan het item, kan het subject het item echter niet juist oplossen, dan is het item dominant aan het subject. Dit kun je laten zien in een geometrisch model door het dominante gegeven rechts van het niet-dominante gegeven te plaatsen.



Enkelvoudige prikkelgegevens zijn niet alleen van toepassing op antwoorden in testitems. Ze komen ook voor in attitudevragenlijsten in de sociale psychologie en bij een medische diagnose.

Individueen zijn te ordenen langs één onderliggende (latente) dimensie. Elk individu verdeelt die dimensie in 2 stukken: één gebied waarin de items liggen die het individu fout heeft beantwoordt en één gebied waarin de items liggen die het individu juist beantwoordt. Verschillende individuen verdelen de dimensie op verschillende plaatsen in tweeën. Dit is informatie op ordinaal niveau.

De items zijn te ordenen langs dezelfde onderliggende (latente) dimensie. Elk item verdeelt de dimensie in 2 stukken: één gebied waarin de individuen liggen die het item fout beantwoorden en één gebied waarin de individuen liggen die het item juist beantwoorden. De verschillende items verdelen de dimensie op verschillende plaatsen in tweeën. Dit is ook informatie op ordinaal niveau.

De scalogram- of simplexstructuur van Guttman is een perfect triangulaire structuur. Hiervoor mag je de items en subjecten door elkaar husselen. Wanneer je de triangulaire structuur hebt bereikt transformeer je dit op de lijn door van links naar rechts de gehusselde items neer te zetten (bijv. c, a, b, e, d) en dan zet je onder de lijn, links van het bijbehorende itempijlje, het aantal subjecten per segment.

Let wel, in de realiteit zijn er ook antwoordpatronen die afwijken van een perfect triangulaire structuur.

	a	b	c	d	e	Fr
1	0	0	1	0	0	20
2	1	1	1	0	0	17
3	0	0	0	0	0	15
4	1	1	1	0	1	15
5	1	0	1	0	0	15
6	1	1	1	1	1	18
7	1	1	1	1	0	35
8	1	0	1	1	0	18
9	0	0	1	1	1	10
10	0	1	1	1	0	37

→

	c	a	d	b	e	Fr
3	0	0	0	0	0	15
1	1	0	0	0	0	20
5	1	1	0	0	0	15
8	1	1	1	0	0	18
7	1	1	1	1	0	35
6	1	1	1	1	1	18
2	1	1	<u>0</u>	1	<u>0</u>	17
4	1	1	<u>0</u>	1	1	15
9	1	<u>0</u>	1	<u>0</u>	1	10
10	1	<u>0</u>	1	1	0	37

Opdat antwoordpatroon 2 zou passen in de triangulaire structuur, zou 11010 moeten veranderd worden in 11110, gelijk aan patroon 7. Dit gaat echter niet, en daarom worden subjecten met patroon 2 in hetzelfde segment geplaatst als subjecten met patroon 7. Wanneer de antwoorden van de subjecten op deze manier worden voorspeld, spreken we van foute reproducties. In bovenstaand voorbeeld maken we bij patroon 2, 1 aanpassing. Dat zijn dus 17 foute reproducties. In totaal krijg je:

- 17 fouten
- 15 fouten
- 2 x 10 fouten
- 37 fouten

Totaal 89 fouten

De *reproductiecoëfficiënt* geeft aan hoe goed het scalogrammodel de oorspronkelijke gegevens verklaart.

$$\text{Rep} = 1 - \frac{\text{aantal fouten}}{\text{aantal antwoorden}}$$

$$\text{Rep} = 1 - \frac{89}{200 \cdot 5} = 0.911 \quad (200 = \text{aantal subjecten}, 5 = \text{aantal items})$$

Soms is er ook een andere dataoplossing mogelijk, de dataoplossing die je dan kiest is de dataoplossing met de grootste reproductiecoëfficiënt.

Hoofdstuk 3. Basisprincipes bij het opzetten van experimenteel onderzoek

Wetenschappelijke kennis wordt in de eerste instantie opgebouwd door uitspraken te doen over de onderlinge samenhang tussen variabelen. Hierbij komen twee problemen kijken. Het eerste is het probleem van de meetbaarheid van psychologische variabelen en het tweede is hoe we de causaliteit uit geobserveerde gegevens kunnen infereren. Hoe kunnen uit gegevens valide conclusies getrokken worden over de onderzoekshypothese?

Er zijn hiervoor drie basisprincipes:

1. het maximaliseren van de experimentele variabiliteit
2. het minimaliseren van de foutenvariabiliteit
3. het maximaliseren van de controle op storende variabelen

3.2. Variabelen en proefopzet

Het skelet van een experiment

Het experiment bevat altijd tenminste één onafhankelijke variabele. Deze variabele wordt gemanipuleerd en de onderzoeker zorgt voor de variatie in deze manipulatie. Deze variatie resulteert in verschillende experimentele condities of behandelingen. Daarnaast is er altijd tenminste één afhankelijke variabele. Dit is de observatievariabele. De bedoeling is na te gaan in welke mate de data afhankelijk zijn van het niveau van de onafhankelijke variabele. De potentieel storende variabelen worden gecontroleerd. Een storende variabele varieert mee met een onafhankelijke variabele en oefent een invloed uit op de afhankelijke variabele.

De bedoeling van een experiment is nagaan in welke mate de data afhankelijk zijn van het niveau van de onafhankelijke variabele. Een onafhankelijke variabele heeft tenminste twee niveaus, maar kan er dus ook meer hebben.

Correlationeel verband / causaal verband

Er is een causaal verband wanneer er gebeurtenissen plaatsvinden als gevolg van andere gebeurtenissen die daaraan vooraf zijn gegaan.

Voorbeeld: Stellen die meerdere malen per week seksuele gemeenschap hebben zonder voorbehoedsmiddelen, krijgen kinderen.

Een correlationeel verband bestaat wanneer er een verhouding van onderlinge afhankelijkheid of beïnvloeding is tussen twee gebeurtenissen.

Voorbeeld: Taiwanese mensen met veel elektrische apparatuur in huis, gebruiken meer anticonceptiemiddelen.

Een onafhankelijke variabele kan als een causale variabele worden gemanipuleerd. Wanneer X dan aanwezig is, wordt er een effect op Y geobserveerd. Wanneer X afwezig is en er geen effect op Y wordt geobserveerd is X een noodzakelijke en voldoende voorwaarde.

Meerdere onafhankelijke en afhankelijke variabelen mogelijk

Een unifactoriële, univariate proefopzet: één onafhankelijke variabele en één afhankelijke variabele.

Een unifactoriële, multivariate proefopzet: één onafhankelijke variabele en meerdere afhankelijke variabelen.

Een multifactoriële, univariate proefopzet: meerdere onafhankelijke variabelen en één afhankelijke variabele.

Een multifactoriële, multivariate proefopzet: meerdere onafhankelijke variabelen en meerdere afhankelijke variabelen.

De implicaties voor de data-analyse:

Bij univariate proefopzetten: vaak een variantie-analyse (ANOVA)

Bij multivariate proefopzetten: vaak multivariate variantie-analyse (MANOVA)

3.3. *Validiteitsvoorwaarden*

Cook & Campbell hebben vier vormen van validiteit opgesteld waaraan experimenteel onderzoek moet voldoen.

1. **Validiteit van statistische conclusie.** Is er een covariatie tussen de onafhankelijke en afhankelijke variabele(n)?
2. **Interne validiteit.** Gegeven dat er een verband is, gaat het om een causaal verband van een geoperationaliseerde variabele naar een andere?
3. **Constructvaliditeit.** Gegeven dat het verband vermoedelijk causaal is, was zijn de theoretische constructen die aan de basis liggen van het verband?
4. **Externe validiteit.** Gegeven dat er vermoedelijk een causaal verband is tussen constructen A en B, kunnen we die relatie dan generaliseren naar andere personen, situaties en momenten en naar andere operationalisaties?

Validiteit van statistische conclusie

Je moet beslissen of de onderzoekshypothese al dan niet aangehouden wordt, dit gebeurt meestal op basis van statistische data-analyse. Onderliggend aan een statistische data-analyse ligt een beslissingsmodel. Dit model legt voorwaarden op aan de gegevens:

- a. Is de studie gevoelig genoeg om iets te concluderen over de covariatie tussen onafhankelijke en afhankelijke variabelen? Dit heeft te maken met power-analyse. Hoe groot moet de steekproef minstens zijn om een effect van een bepaalde grootte te kunnen observeren?
- b. Indien de studie gevoelig genoeg is, is er voldoende evidentie om te besluiten dat de variabelen covariëren? Dit wordt vaak gedaan door de klassieke significantie-toetsen.
- c. Indien er voldoende evidentie is voor covariatie tussen de variabelen, hoe sterk covariëren de variabelen? Dit heeft te maken met de effectgrootte.

Enkele mogelijke bedreigingen van een statistische conclusie validiteit zijn een te lage statistische power, dus bijvoorbeeld een te kleine steekproef, en dat de data niet voldoet aan de voorwaarden van het model, dus dat bijvoorbeeld de variantie-analyse veronderstelt dat de score op de afhankelijke variabele normaal verdeeld is.

Interne validiteit

Een onderzoek is intern valide, als de waargenomen variatie in de afhankelijke variabele op ondubbelzinnige wijze kan toegeschreven worden aan variatie in de onafhankelijke variabele en niet aan variatie in een storende variabele.

Er zijn hierbij twee mogelijke problemen. Het eerste probleem is de richting van het verband ($A \rightarrow B$ of $B \rightarrow A$). Voorbeeld: leidt het kijken naar geweld op TV (A) tot meer agressief

gedrag (B)? Dit kun je onderzoeken, maar de vraag die daarbij gesteld moet worden is ‘kijken agressievere jongeren (B) liever naar gewelddadige Tv-programma’s (A)?’.

Het tweede probleem is de invloed van de storende variabele. Wanneer er twijfel is of de variatie in de afhankelijke variabele te wijten is aan de onafhankelijke variabele of aan een storende variabele is het onderzoek niet intern valide.

Overige termen:

Concomitante variabele (samengaan)

Contaminerende variabele (een variabele die voor verwarring zorgt, besmetting van het experiment)

Alternatieve of rivaliserende hypothese: specificeert het verband tussen de afhankelijke variabele en een andere variabele dan de onafhankelijke variabele

Constructvaliditeit

De interne validiteit heeft te maken met de relatie tussen geoperationaliseerde variabelen. Bij de constructvaliditeit vraagt men zich af of we op basis van de geoperationaliseerde variabelen uitspraken mogen doen over de onderliggende theoretische constructen.

Externe validiteit

Een onderzoek is extern valide, als de conclusies van het onderzoek veralgemeend kunnen worden naar andere participanten, situaties en tijdsmomenten.

Er zijn twee manieren om naar de externe validiteit te kijken. Je hebt de engere betekenis van veralgemening en de bredere betekenis: het belang van de theorie. Bij de eerste is er de veralgemening naar een grotere populatie of naar andere situaties. Bij de tweede wil de onderzoeker vooral theoretische conclusies trekken. Het gevolg van deze theoretische conclusies is dat de externe validiteit hiervan vaak de individuele experimenten overstijgt.

Externe validiteit kan je nagaan via een nieuw onderzoek. Je kunt een directe replicatie uitvoeren, dus de originele studie zo exact mogelijk over doen, of je kunt een systematische replicatie uitvoeren; je doet de originele studie over, maar met systematische variatie wat betreft participanten, situaties, tijdsmomenten en/of operationalisering van variabelen.

Wanneer je de interne validiteit verhoogt door o.a. een striktere experimentele controle op mogelijke storende variabelen, leidt dit vaak tot een specifiekere, artificiële situatie. De vaakgehoorde kritiek hierop is dat je het dan eigenlijk niet meer kan toepassen op het echte leven. Maar de interne validiteit is wel degelijk belangrijk, dit is namelijk de enige manier om te zien welke variabelen daadwerkelijk een inherente correlatie vertonen en of zij dus invloed kunnen hebben op de wereld. Dat is dus ook het belang van de theorie.

3.4. Experimentele vs. foutenvariabiliteit

Psychologen zijn geïnteresseerd in variatie in gedrag en wat die variatie bepaald. Een experimentele hypothese is dat de variatie in de afhankelijke variabele wordt bepaald door een variatie in de onafhankelijke variabele.

Voorbeeld: aangeleerde smaakaversie

De klassieke conditionering: Pavlovs hond.

Vlees = onvoorwaardelijke stimulus (OS)

Saliveren op vlees = onvoorwaardelijke respons (OR)

Toon = voorwaardelijke stimulus (VS)
Saliveren op toon = voorwaardelijke respons (VR)

Voor conditionering:

OS (voedsel) → OR (saliveren)

VS (toon) → oriënteringsreflex

Tijdens conditionering:

VS + OS → OR

Na conditionering:

VS → VR

Totale variantie

$$\sigma^2 = \frac{\sum_{j=1}^p \sum_{i=1}^n (y_{ji} - \overline{Y_{..}})^2}{(np - 1)}$$

- De totale variantie stelt de afwijking van elke score t.o.v. het totale gemiddelde voor.
- Je sommeert dus over n en p
- En je deelt door het aantal proefpersonen -1

n = het aantal subjecten per conditie

p = het aantal condities

Variantie binnen elke groep

$$\sigma_j^2 = \frac{\sum_{i=1}^n (y_{ji} - \overline{Y_{j.}})^2}{(n - 1)}$$

- De variantie binnen elke groep stelt de afwijking van elke score t.o.v. het gemiddelde van de specifieke conditie voor.
- Sommen over het aantal proefpersonen per conditie
- Delen door het aantal proefpersonen per conditie -1

$$\sigma_{BG}^2 = \frac{\sum_{j=1}^p \sum_{i=1}^n (y_{ji} - \overline{Y_{j.}})^2}{p(n - 1)}$$

Gemiddelde kwadraat
binnen de groepen:

GK_{BG}

Variantie tussen groepen

Bij verschillen tussen individuen van verschillende groepen is er één bijkomende bron van variabiliteit en dat is de variabiliteit ten gevolge van manipulatie van een onafhankelijke variabele.

$$\sigma_{TG}^2 = \frac{\sum_{j=1}^p (y_{ji} - \overline{Y..})^2}{(p-1)}$$

Gemiddelde kwadraat tussen de groepen: GK_{Tussen}
--

Vergelijking variantie tussen/binnen groepen

De variantie binnen groepen is een foutenvariabiliteit. De variantie tussen groepen is een foutenvariabiliteit en een experimentele variabiliteit (de variabiliteit ten gevolge van manipulatie van de onafhankelijke variabele).

Wanneer er geen effect uitgaat van de onafhankelijke variabele op de afhankelijke variabele, dan zal $GK_{Tussen} = GK_{binnen}$ (want gaan terug op dezelfde bronnen van variabiliteit, behalve de onafhankelijke variabele). Wanneer er een effect uitgaat van de onafhankelijke variabele op de afhankelijke variabele moet GK_{Tussen} groter zijn dan GK_{binnen} .

3.5. Enkele basisopzetten

De keuze bij een experimentele opzet is beslissen welke participanten welke behandeling zullen krijgen en wanneer. De bedoeling hiervan is om aan zoveel mogelijk validiteitsvoorwaarden te voldoen.

Manipulatie van een variabele tussen subjecten

Elke participant wordt ad random toegewezen aan één groep, dit is dus een volledig gerandomiseerde proefopzet. De participanten van één groep krijgen dezelfde experimentele behandeling.

Manipulatie van meerdere variabelen tussen subjecten

Elke participant wordt ad random toegewezen aan één groep, dit is een volledig gerandomiseerde, factoriele proefopzet.

Manipulatie van een variabele binnen subjecten, met gelijktijdige meting

Elke participant krijgt alle experimentele behandelingen. De eenvoudigste vorm is dat de participant a.h.w. kan kiezen tussen de verschillende niveaus van de onafhankelijke variabele.

Manipulatie van een variabele binnen subjecten met herhaalde metingen

Elke participant krijgt alle experimentele behandelingen. De experimentele condities worden op verschillende tijdsmomenten gerealiseerd.

Manipulatie van meerdere variabele binnen subjecten

???

Manipulatie van één variabele tussen of binnen subjecten en één blokvariabele

De participanten worden in blokken verdeeld, op zo'n manier dat subjecten binnen een blok homogener (minder variabiliteit) zijn dan subjecten tussen blokken. Daarnaast wordt één variabele tussen of binnen subjecten gemanipuleerd, dit is een gerandomiseerde blokopzet.

Voorbeeld: Een test van 4 hoogtemeters in een vliegtuig. De subjecten zijn drie groepen van telkens vier piloten. Groep 1 heeft tussen de 500 en 1500 uur vliegervaring, groep 2 tussen de 1500 en 2500 uur vliegervaring en de 3^e groep heeft 2500 of meer uur vliegervaring. Elke piloot leest 100 keer de hoogte af op elk van de 4 experimentele hoogtemeters (= manipulatie binnen subjecten). De afhankelijke variabele is het aantal fouten dat ze maken.

Gemengde opzetten met manipulatie zowel van variabelen tussen subjecten als van variabelen binnen subjecten

Voorbeeld: Een ergonoom is geïnteresseerd in factoren die de prestatie in een vigilantietaak beïnvloeden. Een vigilantietaak is een taak waarbij er signalen voorkomen met tussenpozen, die onvoorspelbaar en infrequent zijn, zoals een radar of een veiligheidscontrole op de luchthaven.

Experiment: De eerste onafhankelijke variabele is de modaliteit van de signaalpresentatie: is het visueel of auditief. De tweede onafhankelijke variabele is de duur van de "watch": 1, 2, 3 of 4 uur. De afhankelijke variabele is de responsduur t.o.v. het signaal. Er worden 8 participanten ad random verdeeld over 2 groepen van 4 participanten (2x4 opzet). De 2 groepen worden ad random toegewezen aan 2 niveaus van de eerste onafhankelijke variabele.

Deze opzet met minstens één tussen-subject variabele en minstens één binnen-subject variabele noemt men een split-plot proefopzet.

Opmerking: situationele en persoonsvariabelen

In alle voorbeelden, behalve in het voorbeeld van de blokvariabele, was de onafhankelijke variabele steeds een situationele variabele. In principe kan een persoonsvariabele ook een onafhankelijke variabele zijn, zoals intelligentie, geslacht en zelfverzekerdheid.

Er zijn echter enkele verschillen met de situationele variabele. De persoonsvariabele wordt niet gemanipuleerd, omdat dit niet kan. De persoonsvariabele is persoons-eigen. Participanten kunnen hierbij dus niet ad random aan groepen toegewezen worden. In dit geval spreken we van een quasi-experimenteel onderzoek.

Keuze van de proefopzet

Het beslissen van welke participanten welke behandelingen zullen krijgen en wanneer heeft belangrijke implicaties voor:

- a. het maximaliseren van experimentele variabiliteit
- b. het minimaliseren van foutenvariabiliteit
- c. het maximaliseren van controle op storende variabelen

Dit zijn de basisprincipes van experimenteel onderzoek die kans op valide conclusies t.a.v. de onderzoekshypothese moeten verhogen.

3.6. Maximaliseren experimentele variabiliteit

Bij een constante foutenvariabiliteit wordt de kans op significant effect groter naarmate de experimentele variabiliteit toeneemt. Om de onafhankelijke variabele een kans te geven moet je:

- a. Zorgen voor een goede operationalisering van de onafhankelijke variabele, wanneer er geen goede operationalisering is, is de kans op significant effect kleiner.
- b. Indien er keuze is tussen verschillende niveaus van een onafhankelijke variabele, zorg dan dat je voldoende niveaus kiest. Maar bij extreme waarden bestaat de kans dat de constructvaliditeit van de geoperationaliseerde variabele in het gedrang komt. Houdt het daarom wel reëel.
- c. Vermijdt plafond- en vloereffecten. Bij de keuze van niveau(s) van de OV moet je ernaar streven dat de OV aanleiding kan geven tot voldoende variatie in de AV.
Het plafoneffect: OV resulteert in extreem hoge scores op AV (hogere scores zijn niet mogelijk).
Het vloereffect: OV resulteert in extreem lage scores op AV (lagere scores zijn onmogelijke).
- d. Kies een voldoende aantal participanten. Er bestaan procedures die je toelaten te schatten hoeveel participanten minstens aan een bepaald experiment moeten deelnemen om een effect van een bepaalde grootte te kunnen observeren.
- e. De manipulatie check. Dit is een bijkomende observatie om na te gaan of de manipulatie van de OV succesvol was.

3.7. Maximaliseren van de controle op storende variabelen

Een storende variabele maakt geen deel uit van de hypothese en oefent invloed uit op de afhankelijke variabele. Ook varieert ze mee met de onafhankelijke variabele. Een storende variabele kan de interne validiteit van het experiment in het gedrang brengen.

Constantie

Je houdt een potentieel storende variabele constant over alle experimentele condities. Het gevolg hiervan is dat de factor niet meer kan functioneren als een bron van variabiliteit. De geobserveerde variabiliteit in de afhankelijke variabele kan dan niet meer door die storende variabele verklaard worden.

Ook zijn er nog speciale soorten variabelen die constant gehouden moeten worden. Je hebt de *experimentatorvariabelen*. Deze hebben te maken met de persoon die het verloop van het experiment controleert. Een mogelijke oplossing hiervoor is om maar met één proefleider te werken. Maar eenzelfde proefleider is nog geen garantie. De experimentator kan zich ongewild namelijk anders gedragen bij verschillende proefpersonen en op die manier kan uit een invloed uitoefenen op de afhankelijke variabele. Dit is de rol van de verwachtingen van de experimentator.

Ten tweede zijn er nog de verwachtingen van de participanten, ook wel meer bekend als het *placebo-effect*. Dit is een behandeling die op zich geen effect heeft, maar toch tot een effect leidt omdat het subject denkt/gelooft dat de behandeling een effect zal hebben. Placebo-effecten kunnen heel sterk zijn. Mensen kunnen zelfs verslaafd raken aan placebopillen. De oplossing hiervoor is het invoeren van een placeboconditie. Het gevolg hiervan is dat de verwachtingen van participanten in de experimentele en de controleconditie constant zullen zijn en de geobserveerde verschillen in de afhankelijke variabele kunnen dan niet toegeschreven worden aan verschillen in de verwachting.

Wanneer zowel proefleider als de participant conditieblind is noemen we dit een dubbel-blind proefopzet.

Randomisering

Bij Randomisering is de onderliggende veronderstelling dat een potentieel storende variabele, zoals bijvoorbeeld intelligentie, op die manier toch gelijk verdeeld zal zijn over de condities. Dit ligt aan de basis van volledig gerandomiseerde proefopzetten. Het eerste voordeel is dat het de externe validiteit verhoogt. In principe kunnen de bevindingen gegeneraliseerd worden naar alle niveaus van de potentieel storende variabele. Het tweede voordeel is dat je ook voor ongekende storende variabelen kunt controleren, omdat je in principe alle variabelen controleert die samenhangen met de gerandomiseerde entiteiten.

De kans dat randomisering slaagt wordt groter naarmate het aantal subjecten groter is. Let wel op, dit is iets anders dan een random steekproef.

Manipulatie als pseudo-onafhankelijke variabele

Voorbeeld: een onderzoek naar de effectiviteit van twee methodes van rekenonderwijs. Methode 1 is een leerling-gerichte methode, methode 2 is een leerstof-gerichte methode. De onafhankelijke variabele is misschien onder andere afhankelijk van de ervaring van de leerkracht. Vandaar dat de ervaring van de leerkracht een pseudo-onafhankelijke variabele is. Potentieel storende variabelen zou je dan moeten manipuleren alsof het een onafhankelijke variabele zou zijn.

Een speciaal geval is *contrabalancering*. Dit is het manipuleren van een (storende) variabele die te maken heeft met de opzet van het onderzoek zelf. De manipulatie gebeurt tussen subjecten.

- a. Volgorde-effecten bij manipulatie van een OV binnen subjecten: een probleem dat zich zou kunnen voordoen is het carry-over effect. Het feit dat een participant deelnam aan één niveau van een conditie kan een invloed hebben op zijn gedrag in niveaus die later aangeboden worden (vermoeidheid, etc.). Een mogelijke oplossing hiervoor is contrabalancering. Verschillende niveaus van een onafhankelijke variabele worden in een verschillende volgorde aan verschillende (groepen van) proefpersonen aangeboden.
- b. Andere variabelen: bijvoorbeeld de contrabalancering van drukknoppen. De helft van de proefpersonen drukt op de linkerknop bij een verplaatsing en op de rechter bij geen verplaatsing. De andere helft drukt op de rechterknop bij een verplaatsing en op de linker bij geen verplaatsing.
- c. Een speciale vorm van contrabalancering is de Latijns-vierkant-opzet. Dit is ter controle van de volgorde-effecten. Het aantal niveaus van de OV is hierbij groter dan 2 en het aantal participanten is gelijk aan of een veelvoud van het aantal niveaus van de OV. De Latijns-vierkant-opzet gaat terug op een oud raadsel: op hoeveel verschillende manieren kun je n Latijnse letters in een matrix van n kolommen en n rijen rangschikken, zodanig dat elke letter maar één keer in elke rij en in elke kolom voorkomt.

Gelijkstelling (matching)

Voorbeeld: onderzoek naar de effectiviteit van vier methodes van rekenonderwijs. Je kunt de invloed van intelligentie controleren via gelijkstelling. Stel je hebt 20 proefpersonen, deze herschik je allemaal in functie van IQ, in blokken van vier. De vier subjecten met het laagste IQ vormen het eerste blok en de vier subjecten met het hoogste IQ het vijfde blok. Vervolgens wijs je “ad random” de subjecten binnen de blokken toe aan de experimentele condities. Dit houdt in dat bij elke conditie er 5 participanten zijn die allemaal in te delen zijn bij een ander blok van IQ.

De voordelen hiervan zijn dat er min of meer zekerheid is dat de storende variabele in gelijke mate aanwezig is op verschillende niveaus van de OV, dit is analoog aan constantie en daarmee goed voor de interne validiteit. Binnen elke conditie wordt de potentieel storende variabele niet constant gehouden, die is analoog aan randomizing en daarmee goed voor de externe validiteit.

De kern van proefopzetten met een blokvariabele is dat de homogeniteit van de storende variabele groter is binnen de blokken dan tussen de blokken.

De kern van proefopzetten met manipulatie binnen de subjecten, met herhaalde metingen is dat hetzelfde subject binnen hetzelfde blok is en het subject is zijn of haar eigen controle. Het probleem wat je hiermee alleen hebt zijn de carry-over effecten.

Combinatie van methodes

In één specifiek experiment worden verschillende controlemechanismen meestal gecombineerd. Sommige variabelen worden gecontroleerd door de variabele constant te houden, andere variabelen via randomisering, andere variabelen door ze te manipuleren als pseudo-onafhankelijke variabele en andere via gelijkstelling.