

Methodenleer van experimenteel onderzoek

Wetenschappelijk onderzoek is slechts mogelijk doordat systematisch, volgens afgesproken, in de logica verankerde regels tewerk wordt gegaan. Een wetenschappelijk onderzoeker dient gebruik te maken van algemeen aanvaarde onderzoeksmethoden. In de mate waarin hij hiervan afwijkt worden zijn onderzoeksresultaten meer aanvechtbaar. Tevens zal hij de onderzoeksgegevens systematisch en objectief dienen te verzamelen.

In dit hoofdstuk wordt nader ingegaan op de wijze waarop het wetenschappelijk onderzoek kan worden geclassificeerd (naar: Swanborn 1987). Vervolgens worden enkele kernbegrippen behandeld aan de hand van het zogenaamde 'laboratorium experiment'.

1 Classificatie van onderzoek

1.1 Niveau of soort vraagstelling

Voor een systematische behandeling van psychologisch (en ook ergonomisch) onderzoek is het nuttig een onderscheid te maken tussen vier niveaus van onderzoek. Je zou ook kunnen spreken van vier soorten onderzoeksvragen. Deze zijn:

- 1 Beschrijvend onderzoek
- 2 Voorspellend onderzoek
- 3 Verklarend onderzoek
- 4 Toegepast onderzoek

1. Beschrijvend onderzoek

Als eerste kan men een verschijnsel beschrijven. Beschrijvend onderzoek gaat vaak vooraf aan onderzoek op één van de volgende twee niveaus: het is eenvoudig in te zien dat het voorspellen of verklaren van verschijnselen niet goed mogelijk is als men de verschijnselen zelf niet goed kent. Een typerend voorbeeld van beschrijvend onderzoek is het werk van het Centraal Bureau voor de Statistiek (CBS). Zogenaamd verkennend onderzoek of explorerend onderzoek valt ook in de categorie beschrijvend.

2. Voorspellend onderzoek

Een tweede niveau van wetenschappelijk onderzoek betreft het leggen van relaties of het voorspellen van verschijnselen. Een goed voorbeeld daarvan is het onderzoek naar eenhandige bewegingstijd dat beschreven staat in van Leyden (11.5.3).

3. Verklarend onderzoek

Een derde niveau van onderzoek is verklaren, het vaststellen van oorzaak-gevolg relaties. Alleen onderzoek van dit niveau kan in strikte zin leiden tot wetenschappelijke

Methodenleer van experimenteel onderzoek

kennis, waar begrip van de werkelijkheid, uitgedrukt in causale relaties voorop staat. Veel van het in van Leyden geciteerde onderzoek hoort tot deze categorie.

4. Toegepast onderzoek

Tenslotte kunnen we onderscheiden het toegepast onderzoek. Hierbij gaat het in het algemeen om het veranderen van een bestaande situatie. De nadruk ligt niet op het vergaren van -wetenschappelijke- kennis, maar op het verantwoord gebruiken van dergelijke kennis, bijvoorbeeld om de vraag te beantwoorden hoe groot de 'nieuwe rijksdaalder' zou moeten zijn (beschreven in van Leyden, hoofdstuk 1).

1.2 Onderzoeksstrategieën

Behalve verschillende niveaus van onderzoek zijn er ook verschillende onderzoeksstrategieën mogelijk. Met 'onderzoeksstrategie' wordt dan aangeduid hoe het onderzoek wordt opgezet, hoe en onder welke omstandigheden de gegevens worden verzameld. De verschillende strategieën zijn meer of minder passend voor het beantwoorden van een bepaald type onderzoeksvraag.

Relevante strategieën voor de functieleer en ergonomie die we kunnen onderscheiden zijn:

Experiment

Dit wordt in de volgende paragraaf verder uitgewerkt vanwege het grote belang van het experiment voor psychologisch onderzoek.

Enquête

Hieronder wordt verstaan een schriftelijke of mondelinge ondervraging van personen. Deze personen produceren zelf de gegevens. Het begrip 'enquête' moet ruim opgevat worden en betreft bijvoorbeeld alle mogelijke vormen van vragenlijsten, interviews en beoordelingsschalen (rating scales).

Veldonderzoek

Dit wordt gekenmerkt door twee zaken: de onderzoeker vertoeft zelf (kortere of langere tijd) in de onderzochte situatie en de onderzoeker zelf produceert de gegevens, bijvoorbeeld door systematische observaties of het gebruik van checklists. Deze onderzoeksstrategie wordt vooral toegepast als het naturalistische karakter van de onderzochte situatie erg belangrijk is.

Bureauonderzoek

De term zegt het al: het gaat om alle vormen van onderzoek die je van achter je bureau kunt uitvoeren. Ook deze onderzoeksstrategie dekt een groot aantal datavergaringsmethoden, zoals analyse van statistische gegevens, secundaire analyse van door andere onderzoekers verzameld materiaal, inhoudsanalyse van documenten, maar ook literatuuronderzoek.

Simulatie

De onderzoekssituatie is dan een nabootsing van relevante aspecten uit de werkelijkheid, waarover de onderzoeker uitspraken wil doen. Voorbeelden van een vrijwel volledige nabootsing van de werkelijkheid zijn de scheepsbrug-simulator en de vliegsimulator. Simulaties van een klein gedeelte van de werkelijkheid worden bijvoorbeeld gebruikt bij het bestuderen van het gedrag van procesoperators. De simulatie wordt vrijwel altijd in combinatie gebruikt met een andere strategie. Vaak is dit het experiment, om de eenvoudige reden dat een experiment in de werkelijke situatie niet mogelijk is, bijvoorbeeld als de onderzoeker het gedrag van procesoperators wil bestuderen in relatie tot het optreden van ernstige procesafwijkingen. Een mock-up kan ook beschouwd worden als een simulatie.

2 Het laboratorium experiment

2.1 Definitie van een experiment

Zimney (1961) geeft als definitie van een experiment: Een objectieve observatie van verschijnselen die tot stand worden gebracht in een streng gecontroleerde situatie waarin een of meer factoren worden gevarieerd en de andere constant worden gehouden.

De definitie bevat een aantal elementen die we nader zullen beschouwen. De eis van **objectieve observatie** impliceert dat de experimentator geen neiging mag vertonen om vooringenomen, partijdig waar te nemen. Dit zou kunnen vóórkomen omdat hij de theorie die getoetst wordt graag bevestigd wil zien. Uit onderzoeken van Rosenthal (1966) is komen vast te staan dat attitudes en verwachtingen inderdaad de uitkomsten van het experiment kunnen beïnvloeden. Omdat in experimenten de uitkomsten bestaan uit metingen van de een of andere soort, dient deze meting accuraat, objectief te geschieden¹.

Onder **verschijnselen** verstaan we binnen een psychologisch experiment elk waarneembaar gedrag zoals handelingen, verbale uitingen en antwoorden op vragenlijsten. Niet waarneembare interne processen en toestanden van het individu worden hierbij uitgesloten; verbale rapportering door proefpersonen van hun interne processen (introspectie) en fysiologische indicatoren van deze processen komen echter wel voor beschouwing in aanmerking; ze vormen immers waarneembaar gedrag.

In een experiment wacht de onderzoeker niet af tot de verschijnselen plaatsgrijpen, maar hij beïnvloedt het ontstaan daarvan. Vandaar dat de definitie spreekt over 'phenomena which are made to occur' (verschijnselen die tot stand worden gebracht). De onderzoeker manipuleert de verschijnselen, laat ze gebeuren. Dit doet hij door de omstandigheden te manipuleren. De omstandigheden die door de onderzoeker worden aangebracht of gevarieerd noemt men de **onafhankelijke variabelen**. De

¹ De eis van objectiviteit is niet altijd realiseerbaar. De eis dient soms te worden afgezwakt tot 'intersubjectieve overeenstemming tussen onderzoekers'. De onderzoekers dienen daarbij te streven naar onbevooroordeeldheid.

Een onderzoeker is geïnteresseerd in het effect van omgevingstemperatuur op de leerprestatie. Hij kiest een bepaalde leertaak en laat deze door drie groepen proefpersonen leren; elke groep bij een andere temperatuur, te weten: 15°, 20° en 25°. In dit geval is de hoogte van de temperatuur de onafhankelijke variabele. Deze wordt door de onderzoeker gemanipuleerd. De onafhankelijke variabele neemt één of meerdere waarden aan. Er kunnen meerdere onafhankelijke variabelen in een experiment aangebracht worden. Het maximale aantal experimentele **condities** is dan het product van het aantal onafhankelijke variabelen en waarden die deze aan kunnen nemen. De leerprestatie (die afhangt van de gemanipuleerde variabele) is hier de afhankelijke variabele.

effecten worden vastgesteld op grond van bepaalde maten, de **afhankelijke variabelen** genaamd.

Het laatste deel van de definitie luidt: 'waarin een of meer factoren worden gevarieerd en de andere constant worden gehouden'. Er is dus zowel **variatie** als **constantie**. Alleen de door de onderzoeker bewust beïnvloede variabelen (de onafhankelijke variabelen) mogen variëren. Deze dienen de verschillen in uitkomsten op de afhankelijke variabelen te veroorzaken en daarmee te verklaren. Dat wil zeggen dat deze verschillen uitsluitend mogen worden toegeschreven aan variatie in de onafhankelijke variabelen, niet aan andere variabelen.

Stel dat in het genoemde leerexperiment zou blijken dat de proefpersonen (ppn.) in de groep die in de 15°-conditie leert een betere prestatie leveren dan de ppn. uit de 25°-groep; dan lijkt de conclusie voor de hand te liggen dat de hogere temperatuur oorzaak is van een slechtere leerprestatie. Maar als de ppn. in de 15°-groep nu eens intelligenter of gemotiveerder zijn dan die in de 25°-groep dan kunnen de verschillen in leerprestatie ook daardoor verklaard worden. De verschillen kunnen dan niet eenduidig worden toegeschreven aan de onafhankelijke variabele. Daarom dient er voor gezorgd te worden dat de variabelen die óók verschillen op de afhankelijke variabele zouden kunnen veroorzaken (zogenaamde **secondaire variabelen**) in de verschillende condities in gelijke mate invloed uitoefenen.

2.2. Keuze van het design

Een andere mogelijkheid om het leeronderzoek waarvan hierboven sprake was op te zetten is het aanbieden van een aantal leertaken aan *dezelfde* ppn; de ene leertaak wordt dan geleerd onder 25°, de tweede onder 20° en de derde onder 15°. Dan heeft variatie over de condities *binnen* de ppn. plaats.

De vraag welke opzet van het onderzoek het best passend is voor de beantwoording van een vraagstelling is een kwestie van designkeuze. De term **design** betekent opzet, plan of strategie van onderzoek².

Met het zojuist genoemde voorbeeld van een experiment is een belangrijke tweedeling van onderzoeksopzetten (designs) gegeven. De eerstgenoemde mogelijkheid om de

² Omdat de term design ook in het spraakgebruik van Nederlandse onderzoekers is ingeburgerd - de term 'opzet' of 'ontwerp' wordt vrijwel nergens gebruikt - zullen we verder van design spreken.

Functieleer

temperatuurinvloed op de leerprestatie te onderzoeken is om voor elke temperatuur (elke conditie) een aparte groep ppn. te nemen. Er zijn dan drie aparte, separate groepen ppn. De (gemiddelde) scores van de drie groepen kunnen dan worden vergeleken. We hebben dan een zgn. **aparte groepen design**.

De tweede mogelijkheid, waarbij de verschillende condities binnen dezelfde pp. worden gevarieerd en waarbij de (gemiddelde) scores op elk van die condities worden vergeleken, wordt een **enkele-groep design** (ook wel counterbalanced design) genoemd.

De secundaire (relevante) variabelen worden in de verschillende designtypen op verschillende wijzen onder controle gehouden. Teneinde dit verder na te gaan moeten we eerst de verschillende typen relevante variabelen onderscheiden.

2.3 Relevante (secondaire) variabelen

Onder relevante variabelen verstaan we die variabelen waarvan bekend is dat ze de afhankelijke variabelen beïnvloeden of waarvan vermoed wordt dat ze dit doen. De laatste soort variabelen worden wel **potentiële relevante variabelen** genoemd. Beide typen variabelen dienen veiligheidshalve onder controle te worden gehouden. We spreken over beide soorten verder als **relevante variabelen**.

Er worden drie categorieën relevante variabelen onderscheiden, te weten:

- 1 Persoonsvariabelen
- 2 Situationele variabelen
 - omgevingsvariabelen
 - taakvariabelen
- 3 Positie-variabelen

1. Persoonsvariabelen

Bij persoonsvariabelen kan worden gedacht aan intelligentie, sekse, motivatie, leeftijd, enzovoort. In het voorbeeld hierboven hebben we intelligentie- en motivatieverschillen tussen de ppn. genoemd als mogelijke oorzaak van het gevonden verschil in leerprestatie onder de verschillende temperaturen.

2. Situationele variabelen

Deze vallen uiteen in (a) omgevingsvariabelen en (b) taakvariabelen.

Onder omgevingsvariabelen wordt dan verstaan: de vochtigheidsgraad, lawaai, licht, temperatuur, tijd van de dag, tijd van de werk -of schoolweek, enzovoort. Ook proefleider-effecten vallen hieronder.

In het hierboven beschreven onderzoek naar het effect van de temperatuur werd de temperatuur bewust gevarieerd (vormt een onafhankelijke variabele). Andere omgevingsvariabelen zoals vochtigheidsgraad en lawaai dienen onder controle te worden gehouden.

Bij taakvariabelen kan gedacht worden aan de moeilijkheidsgraad van de taak. Als ppn. een aantal taken krijgen onder verschillende condities (bijv. leertaken onder verschillende temperaturen) dan mogen die taken niet verschillen in moeilijkheidsgraad.

3. Positie-variabelen

Als ppn. kort achter elkaar een aantal taken krijgen uit te voeren (onder verschillende condities) dan treedt er *gewenning* op aan het soort taken; ook *oefenings*-, *vermoeidheids*- en *vervelingseffecten* kunnen een rol spelen. Afhankelijk van de positie van de taak (conditie) in de sequentie van taken kan de ene conditie sterker door deze sequentie-effecten worden beïnvloed dan de andere. De eerdere taken profiteren minder van oefeningseffecten, de laatste worden sterker door vermoeidheid en verveling beïnvloed. De positie van de condities in de volgorde dient zodanig te zijn dat deze positie-effecten in de verschillende condities een (ongeveer) even grote rol spelen.

Bij taken die herhaald, maar met grotere tussentijd (in de orde van weken of maanden) worden uitgevoerd kunnen *groei*- en *rijpingsprocessen* een rol gaan spelen. Een verbetering in prestatie hoeft dan niet (uitsluitend) het gevolg te zijn van de experimentele condities.

Nu we een nader onderscheid binnen de relevante variabelen hebben gemaakt willen we nagaan in welke mate het effect van elk der drie soorten relevante variabelen in het aparte-groepen design en in het enkele-groep design wordt geneutraliseerd.

2.4 Aparte groepen designs

De aparte-groepen designs worden verdeeld in twee subtypen:

1. Gematchte groepen designs
2. Gerandomiseerde groepen designs

1. Gematchte groepen designs

Bij dit designtype worden de persoonsvariabelen onder controle gehouden door deze constant te houden of deze op te nemen in het design. Als bijvoorbeeld bekend is dat intelligentie een relevante variabele is, dan kan gezorgd worden dat de ppn. in de verschillende groepen een zelfde intelligentie hebben (bijvoorbeeld een intelligentie quotiënt (IQ) tussen 110 en 115). Deze methode van constant houden heeft echter twee nadelen. Om te beginnen wordt het beschikbare aantal (vaak schaarse) ppn. daardoor erg beperkt, terwijl verder op grond van de resultaten slechts beperkte uitspraken mogelijk zijn. Generalisering van de resultaten over ppn. met hoger of lager IQ dan de onderzochte zijn niet verantwoord.

Het kan verstandig zijn om de intelligentie op een andere manier onder controle te houden. Dit kan door deze over een zo breed mogelijke 'range' te laten variëren, maar de invloed toch in elke conditie (nagenoeg) even groot te doen zijn. Daarbij zijn in het algemeen de volgende stappen nodig:

(toegepast op het zojuist genoemde voorbeeld)

- 1 Bepaal van alle beschikbare ppn de IQ-score op de relevante variabele.
- 2 Plaats alle beschikbare ppn. in volgorde van hun IQ-score.

Functieleer

- 3 Maak, uitgaande van de ppn. met de laagste score, blokken van ppn. van een even groot aantal als er condities zijn.
- 4 Wijs de proefpersonen in elk blok volgens toeval (random) toe aan een van de condities.

In Tabel 1 en 2 is een voorbeeld van deze procedure te vinden voor het

pp.	IQ	pp.	IQ	pp.	IQ	pp.	IQ	pp.	IQ
1	97	4	103	7	111	10	115	13	121
2	99	5	105	8	113	11	116	14	121
3	99	6	108	9	113	12	120	15	124

tabel 1 Proefpersonen in het temperatuureffect leerexperiment ingedeeld in blokken op grond van intelligentievolgorde

temperatuureffect leerexperiment. De ppn. worden, omdat er drie condities zijn, opgedeeld in blokken van drie. Binnen die blokken staan de ppn. in toenemende scorevolgorde (Tabel 1).

Zouden de ppn. in elk blok in rangordevolgorde aan de condities worden toegewezen, dan zou het gemiddelde IQ in conditie 3 hoger zijn dan conditie 2 en daar weer hoger dan in conditie 1. De gemiddelden zouden respectievelijk zijn: 109.4, 110.8 en 112.8. Om deze reden wordt de toewijzing van de ppn. aan de condities volgens een rondom procedure binnen de blokken bepaald. Daarom spreken we van de **methode van de gerandomiseerde blokken** (randomized blocks method).

Door werking van het toeval zijn blijkens tabel 2 de intelligentie-invloeden nu beter onder controle gehouden. De gemiddelde IQ's zijn nu respectievelijk: 111.4, 110.4 en 111.2.

groep 1 15°		groep 2 20°		groep 3 25°	
pp.	IQ	pp.	IQ	pp.	IQ
3	99	1	97	2	99
4	103	6	105	5	108
7	111	8	113	9	113
12	120	11	116	10	115
15	124	13	121	14	121
Gemiddelde :					
111.4		110.4		111.2	

tabel 2 Proefpersonen ingedeeld in verschillende temperatuurcondities door matching volgens de methode van de gerandomiseerde blokken

Er geldt overigens dat het moeilijk wordt om te matchen als er sprake is van twee of meer relevante variabelen.

2. Gerandomiseerde groepen designs

Het verschil van dit designtype met de gematchte groepen designs is dat bij de gerandomiseerde groepen designs de controle van relevante variabelen volledig aan

Methodenleer van experimenteel onderzoek

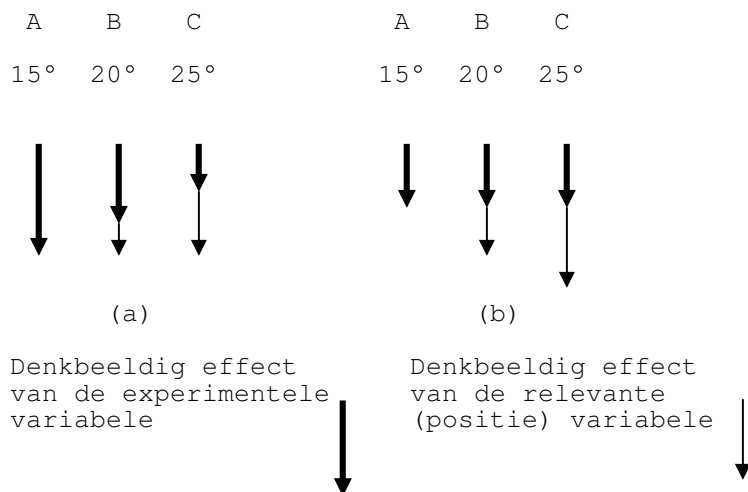
het toeval wordt overgelaten. Deze werkwijze is gebaseerd op de - met name voor grote groepen ppn. gerechtvaardigde - verwachting dat de relevante variabelen in de verschillende condities in dezelfde mate invloed zullen uitoefenen op grond van de randomiserings-procedure. Voor het temperatuureffect onderzoek, zoals dat hierboven werd geschetst, zou dat inhouden dat de ppn. - zonder de IQ-score in aanmerking te nemen - op grond van toeval worden toegewezen aan de condities. Dit in de verwachting dat de gemiddelde IQ-score in de verschillende condities nagenoeg gelijk zal zijn.

Gerandomiseerde groepen designs worden gebruikt in plaats van gematchte groepen designs als er geen relevante variabelen bekend zijn, of als de meting hiervan kostbaar of tijdrovend is, of als matching op relevante variabelen moeilijkheden oplevert (als er bijvoorbeeld veel relevante variabelen zijn of als extreme scores voorkomen; zoals hiervoor genoemd).

2.5 Enkele-groep designs (uitgebalanceerde designs)

Kenmerkend voor het separate-groepen design is dat elke pp. *slechts één* conditie krijgt aangeboden. Kenmerkend voor het enkele-groep design is dat elke pp. *alle* condities krijgt aangeboden.

Als er bijvoorbeeld sprake is van drie condities (A, B en C) dan zouden deze condities kunnen worden aangeboden in de volgorde A, B, C. Toegepast op ons voorbeeld-experiment zou achtereenvolgens geleerd worden onder 15°, 20° en 25°. Op grond van oefeningseffecten zou de prestatie (los van echte temperatuureffecten) bij de laatste leertaak (onder 25°) beter kunnen verlopen dan onder de eerste en de tweede condities. In figuur 1 zijn de denkbeeldige effecten in beeld gebracht.



Figuur 1

Denkbeeldige effecten van experimentele en relevante (positie) variabelen in het temperatuureffect leerexperiment

De effecten zijn denkbeeldig (hypothetisch) omdat de gescheiden effecten van

Functieleer

experimentele en relevante variabelen niet apart worden gemeten. Wat wel wordt gemeten is het effect van beide samen (de totale pijllengte). Splitsing in twee soorten invloeden (van experimentele en van niet-experimentele relevante variabelen) heeft slechts plaats ten behoeve van de lezer, om de (methodo)logische redenering over de controle van de relevante variabelen op te bouwen.

We zouden bij het type design volgens figuur 1.a geen verschil vinden. Het verschil in effect van de experimentele variabele zou verdoezeld worden door het ongelijke, niet geneutraliseerde *positie-effect*. Anders gezegd: we vinden *ten onrechte geen verschil*. Het is ook mogelijk dat we door onvoldoende controle van relevante variabelen *ten onrechte verschil* vinden.

Deze situatie vinden we afgebeeld in figuur 1.b.

Hoe kunnen we dit soort positie-effecten voorkomen? De oplossing die hiervoor wordt gekozen is om de effecten van een bepaalde conditievolvergorden te neutraliseren door deze vergorden (1) òf te laten volgen door een andere vergorden òf (2) andere ppn. een andere conditievolvergorden aan te bieden. Daardoor ontstaat voor elke conditievolvergorden een tegenwicht waardoor de effecten in balans blijven. We spreken om deze reden van **uitbalanceren** (counterbalancing). In het eerstgenoemde geval worden de positie-effecten gecontroleerd *binnen iedere pp.*; in het andere geval heeft deze controle plaats *binnen een groep ppn.* We spreken om deze reden van (1) binnen-proefpersoon uitgebalanceerde designs en (2) binnen-groep uitgebalanceerde designs.

2.6 Controle van de relevante variabelen

Ten aanzien van de *persoonsvariabelen* geldt dat deze in het ene design-type veel meer 'per definitie' onder controle worden gehouden dan in het andere. Omdat binnen de uitgebalanceerde designs alle condities aan dezelfde pp. worden aangeboden zijn de persoonsvariabelen op al deze condities van even grote invloed. Controle van persoonsvariabelen levert in de uitgebalanceerde designs dan ook per definitie geen problemen.

In de aparte-groepen designs daarentegen dient er gewaakt te worden tegen een verschillend effect van deze variabelen in de verschillende condities (een zogenaamd **differentieel effect**). Door matching van relevante variabelen en door randomisering kan dit worden voorkomen.

Voor de *situationele variabelen* geldt met de nodige wijzigingen hetzelfde als voor de persoonsvariabelen. Deze worden over het algemeen door constant houden onder controle gehouden. Zo worden in bijvoorbeeld een leerexperiment het aantal te leren woorden in een lijst, het tijdsinterval tussen de trials, de temperatuur, de belichting, de tijd van de dag, en dergelijke constant gehouden over de verschillende condities.

In een uitgebalanceerd design, waar elke pp. alle condities krijgt aangeboden, is de controle van situationele variabelen nauwelijks problematisch. In de designs waarbinnen met separate groepen wordt gewerkt kan het al gauw voorkomen dat de omstandigheden in de ene groep niet gelijk zijn aan die in de andere. Bewuste controle is hier een vereiste.

De *positiefvariabelen* vormen daarentegen in de aparte-groepen designs nauwelijks een probleem. Omdat elke conditie aan een aparte groep wordt aangeboden wegen

Methodenleer van experimenteel onderzoek

'positie'-effecten (bedoeld worden effecten van herhaalde trials binnen een conditie, zoals oefenings- en vermoeidheidseffecten) in elke groep en daarmee in elke conditie, even zwaar.

Positie-effecten dienen binnen de uitgebalanceerde designs een constante bron van zorg te zijn. Steeds moet worden nagegaan of een goede uitbalancering wel verzekerd is. Dit is ondermeer afhankelijk van het al dan niet lineair zijn van de positie-effecten.

2.7 Factoren die de designkeuze beïnvloeden

Uitgebreidheid (tijdrovendheid) van de taken die de condities representeren. Indien de taken veel pp-tijd in beslag nemen is een binnenproefpersoon design niet wenselijk. Een binnen-groep design, waarin alle ppn. alle condities een keer krijgen aangeboden is al minder tijdrovend. Een aparte-groepen design is het minst tijdrovend per pp.

Herhaalde taakuitvoering. Als het niet mogelijk is om een taak meer dan eens te laten uitvoeren door dezelfde ppn. (denk aan positie-effecten), dan is een binnen-proefpersoon design ongeschikt. Een binnen-groep- of een aparte-groepen design is adequater.

Differentiële uitval van proefpersonen. Dit doet zich voor als deelname aan de ene conditie bijvoorbeeld aantrekkelijker is dan die aan de andere conditie. Alleen de echte doorzetters blijven dan aan de minder aantrekkelijke conditie deelnemen. Daardoor zijn de ppn. in de verschillende condities niet meer vergelijkbaar. Als een aparte-groepen design wordt gebruikt ontstaat dus een vertekend beeld van het conditie-effect. Als een dergelijke differentiële uitval te verwachten valt dan dient een binnenproefpersoon design te worden overwogen. In dat geval valt voor elke conditie één pp. (dezelfde) uit.

Aantal beschikbare proefpersonen. De binnen-proefpersoon designs vereisen weinig ppn. In principe is een pp. voldoende voor het uitvoeren van het experiment. De binnen-groep designs vereisen meer ppn. Afhankelijk van het type binnen-groep design varieert dit aantal. Het aantal kan groot worden als er veel condities zijn.

Opmerking

Natuurlijk zijn de hierboven genoemde factoren ondergeschikt aan de methodologische eisen die aan het design gesteld worden. In de praktijk functioneren de hierboven genoemde factoren vaak als een net van randvoorwaarden waarbinnen het design ontworpen moet worden.

2.8 Factoriële designs

Om de invloed van twee of meer onafhankelijke variabelen op de afhankelijke variabele vast te stellen gebruiken onderzoekers een **factorieel design**. Binnen dit factoriële design worden twee of meer variabelen tegelijkertijd en systematisch gevarieerd.

Functieleer

	Man	Vrouw	
10 min.	3.4	1.6	2.5
20 min.	4.2	3.6	3.9
	3.8	2.6	

Tabel 3. De invloed van de sekse van de communicator en de duur van de boodschap op de attitudeverandering.

In tabel 3 is het design en de denkbeeldige uitkomsten geschetst van een onderzoek naar attitudeverandering. De onafhankelijke variabelen zijn de sekse van de communicator en de duur van de boodschap. Er is dus sprake van twee onafhankelijke variabelen (factoren) met elk twee niveaus. Dit is het eenvoudigste factoriële design, een **2 x 2 design** genaamd. De afhankelijke variabele is de mate van attitudeverandering in positieve richting. De gemiddelde verandering is in de cellen van het design weergegeven. Aan de randen staan de effecten weergegeven per niveau van de onafhankelijke variabele. Daarbij is afgezien van de invloed van de tweede onafhankelijke variabele; dit zijn de hoofdeffecten. Uit de tabel blijkt dat als we de sekse van de communicator niet mede in beschouwing nemen, de attitudeverandering groter is na een 20 minuten-boodschap (3.9) dan na een 10-minuten boodschap (2.5). Er is dus sprake van een hoofdeffect van de factor 'duur van de boodschap'. Er is eveneens een hoofdeffect van de factor 'sekse' (3.8 versus 2.6).

Hoe is nu uit de tabel na te gaan of er sprake is van **interactie**? Dit is vast te stellen door de verschillen tussen de beide celrijen of de beide celkolommen te vergelijken. Als we dit voor de rijen doen, dan blijkt dat de invloed van de sekse van de communicator groter is bij een 10 minuten-boodschap (verschil: $3.4 - 1.6 = 1.8$) dan bij een 20 minuten-boodschap (verschil: $4.2 - 3.6 = 0.6$). Het effect van de ene onafhankelijke variabele (sekse) wordt dus mede bepaald door het effect van de tweede onafhankelijke variabele (duur van de boodschap). In dat geval spreken we van *interactie*.

Of de gevonden hoofd- en interactie effecten significant zijn dient door statistische toetsing te worden bepaald. Variatieanalyse is hiervoor de geëigende statistische analyse methode. Het tweefactor design (bijv. het 2 x 2 design) kan worden uitgebreid met een derde factor zodat er bijv. een 2 x 2 x 3 design ontstaat bestaande uit 12 aparte condities.

Kerlinger (1973, p.257) vermeldt onder andere de volgende voordelen van factoriële designs:

- 1 Omdat er meerdere onafhankelijke variabelen worden gemanipuleerd kunnen meerdere hypothesen tegelijkertijd worden getoetst.
- 2 Als er sprake is van een storende variabele (**confounding variabele**) dan kan deze onder controle worden gebracht door opname hiervan in het design. Deze storende variabele wordt dan één van de factoren die systematisch wordt meegevariëerd. Nevenvoordeel is dat de invloed en significantie van deze variabele mede kan worden bepaald. Als bijvoorbeeld sekse een storende variabele is dan kunnen uitsluitend ppn. van één sekse worden opgenomen (controle door constant houden). Het is echter ook mogelijk om de sekse als een factor te beschouwen.

Methodenleer van experimenteel onderzoek

Hoofdeffecten van en interactie effecten met sekse kunnen dan worden berekend.

- 3 Er kunnen hypothesen over interacties worden getoetst. Het toetsen van hoofdeffecten vereist geen factorieel design, maar voor het toetsen van interactie effecten is dat type design onontbeerlijk.

Een nadeel dat in bepaalde gevallen aan factoriële designs is verbonden, is dat als er sprake is van een complex factorieel design, het design veel cellen bevat en daardoor zijn erg veel proefpersonen nodig. Zo bevat bijv. een $2 \times 3 \times 4 \times 2$ design 48 cellen. Als er per cel 10 ppn. nodig zijn dan zijn er 480 ppn. nodig. Dit kan op problemen stuiten. Als ppn. ruim voorhanden zijn en het geheel niet te tijdrovend wordt, hoeft dit echter geen bezwaar te zijn.

De theorie uit paragraaf 2 wordt meer uitgebreid behandeld in het boek 'Methoden van psychologisch onderzoek' (de Leeuw, Pelle & van der Veer 1983).

Literatuur

F.N. Kerlinger, E.J. Pedhazur, *Multiple regression in behavioral research*, Holt Rinehart & Winston, New York, 1973

L. de Leeuw, P.G. Pelle, G.C. van der Veer, 'Methoden van psychologisch onderzoek; het verzamelen van data; statistiek', In: J.F. Orlebeke, P.J.D. Drenth, R.H.C. Jansen, C. Sanders, *Compendium van de psychologie, dl. 8*, Coutinho, Muiderberg, 1983

dr. J. van Leyden Sr, *Psychologische Functieleer*, Bohn Stafleu Van Loghum, Houten/Antwerpen, 1989

R. Rosenthal, *Experimentor effects in behavioral research*, Appleton-Century-Croft, New York, 1966

P.G. Swanborn, *Methoden van sociaal wetenschappelijk onderzoek*, Boom, Meppel, 1987

G.H. Zimney, *Method in experimental psychology*, Ronald Press, New York, 1961