

Hoorcollege 2

hoofdstuk 7

Het experiment

Bij een experiment in de gedragswetenschappen is er een vergelijking van proefgroepen, het liefst inclusief een controlegroep. Dingen en waarnemingen zijn onzeker, daardoor is een controlegroep belangrijk. Experimenten zijn vooral sterk in het toetsen van causale hypothesen. Om causaliteit aan te tonen zijn 3 voorwaarden die aanwezig moeten zijn:

- Eerste voorwaarde: associatie betekent nooit dat er een causaal verband is. Er moet een relatie tussen oorzaak en gevolg te vinden zijn. De oorzaak is dan de onafhankelijke variabele en het gevolg de afhankelijke.
- Tweede voorwaarde: de oorzaak komt voor het vervolg in tijd.
- Derde voorwaarde: er mogen geen alternatieve verklaringen zijn voor de relatie. Er kan dus geen gemeenschappelijke oorzaak zijn.

Om causaliteit te bewijzen gebruiken exacte wetenschappers vaak het experiment. Het experiment heeft voordelen en nadelen. Voordelen zijn dat het een duidelijke en simpele logica is, daarnaast isoleert het een oorzakelijke verband. Het focust wel maar op beperkt aantal variabelen en heeft wel praktische en ethische aspecten bij mensen.

Onzuivere experimenten wordt het pre experimenteel genoemd. Daartegenover staat echt experimenteel en het middendeel heet quasi experimenteel.

Vervolgens kunnen experimenten ook ingedeeld worden in independent groups designs (dit komt steeds meer voor naarmate een experiment meer echt experimenteel wordt, de groepen van participanten zijn dan onafhankelijk). Dit kan worden berekend met de independent sample t-test. Repeated measures designs (dit kan worden berekend met een paired-sample t-test).

Pre-experimentele designs

Pre-experimentele onderzoeksopzetten missen belangrijke onderdelen van een goede experimentele opzet. Er zijn drie verschillende pre-experimentele onderzoeksopzetten.

- Een one-shot case study: één groep met alleen een voormeting en zonder een controlegroep. De groep proefpersonen wordt experimenteel behandeld.
- Een one group pretest-posttest design: een voormeting en nameting bij een groep. Voor iedere persoon wordt dus dezelfde meting gedaan. Het gevaar bij dit experiment is dat history een rol kan spelen, evenals maturation (innerlijke factoren die zorgen voor een verandering van een persoon, bijvoorbeeld als proefpersonen honger krijgen en het gedrag daardoor verandert) en testing (proefpersonen leren de test kennen en kunnen dan wenselijk antwoorden).
- Een static group comparison: hierbij worden twee groepen vergeleken onder twee condities. Je hebt dan een experimentele conditie en een controle conditie. Dit is een independent group design. Omdat deze groepen verschillend van elkaar vergeleken worden gebruiken we dus de independent sample t-test om het te berekenen. Je kan hier te maken krijgen met een selection bias. Dit is hoeveel toe te wijzen is aan toeval. Je kan er dus minder je best doen. Een ander probleem is mortality, mensen die uitvallen met het onderzoek, je mist dan een aantal metingen. Een oplossing is om te matchen. Je

koppelt dan twee mensen uit verschillende groepen. Deze mensen ga je met elkaar vergelijken uit de verschillende groepen. Je probeert mensen te vinden die erg op elkaar lijken.

Om te kijken aan welke kant van de trechter de experimenten zitten kunnen de trechters in de PowerPoint van methode hoorcollege 1 bekeken worden.

Quasi-experimentele designs

Quasi-experimentele onderzoeksopzet: mensen vullen maandelijks testjes in en halverwege is er een manipulatie. Er zijn twee verschillende designs:

- Two-group pretest-posttest design: het is combinatie van de een independent groups en repeated measures designs, er is dus sprake van twee groepen en er zijn voormetingen en nametingen. Manipulatie vindt plaats tussen de voor- en nameting. Voor en na de manipulatie is er een experimentele en een controlegroep. In de PowerPoint is hier een schematisch te zien hoe de verbanden eruitzien. Het probleem van de selection bias en het mortality effect wordt opgelost bij de voormeting maar bij de nameting is het nog hetzelfde. Daarnaast kunnen voor de twee groepen het testing, history en maturation effect niet generaliseerd worden want misschien hebben zij wel andere factoren beleefd tijdens het onderzoek.
- Interrupted time series (halverwege de therapie een manipulatie). Deze methode bevat meerdere voor- en nametingen. Dit zorgt voor het oplossen van het probleem van testing, history en het maturation effect.

Echt experimentele designs

- Randomized two-group posttest only design: de units worden willekeurig toegewezen aan de condities. Iedereen is dan aan elkaar gelijk omdat het random gekozen wordt. Ze gaan vanuit het independent group design. De verschillen tussen de groepen zijn dan toeval.
- Klassiek experimenteel design: er is dan een voor en een na meting bij twee gerandomiseerde groepen. Het kan fout gaan door toeval maar daarom is de voormeting belangrijk. Dan weten we al hoe de groepen ten opzichte van elkaar staan. Er blijft een controle en een experimentele conditie.
- Solomon four-group design: er zijn vier groepen.
De eerste groep heeft een voormeting, een behandeling en een nameting.
De tweede groep heeft een voormeting, geen behandeling en een nameting.
De derde groep heeft geen voormeting, wel een behandeling en een nameting.
De vierde groep heeft geen voormeting, geen behandeling maar wel een nameting. Hierdoor kan het test effect opgelost worden omdat er geen voormeting is.

Designs met meer dan 2 condities

Er zijn designs die meer dan twee condities hebben bij experimenten.

De eerste zijn independent group designs (bijvoorbeeld zware oefening, lichte oefening, geen oefening).

Factorieel design: als er meerdere onafhankelijke variabelen zijn in het onderzoek moeten ze allemaal meegenomen worden in het experiment. Er moet dan gekruist worden om alle verschillende condities weer te geven. Daarmee kunnen interacties tussen onafhankelijke variabelen onderzocht worden.

Voorbeeld van een factorieel experiment (twee variabelen manipuleren): “Wie mooi is heeft talent?”

De opzet is als volgt.

- Mannelijke studenten krijgen de opdracht om een cijfer (schoolcijfer van 1-10) te geven aan een essay zogenaamd van een ‘studente’. Bij het essay zat een foto van die studente. Random kreeg 1/3 van de proefpersonen een ‘fysiek aantrekkelijke’ foto, 1/3 een onaantrekkelijke en 1/3 een neutrale foto. Random kreeg de helft van de beoordelaars een slecht essay, de andere helft een goed geschreven essay. (Essays in werkelijkheid geschreven door onderzoekers).
- Er is sprake van drie variabelen: onafhankelijk: gemanipuleerd (mooiheid en kwaliteit essay) en de afhankelijke variabele: de beoordeling.
- Er waren 3 hypothesen: mooiheid heeft invloed op de beoordeling van het essay, kwaliteit heeft invloed op de beoordeling van het essay, mooiheid en de kwaliteit van het essay hebben een gezamenlijke invloed op de beoordeling van het essay (indirect). Mogelijke resultaten: oordelen zijn positief als schrijfster knap was. Schrijfster van goede essays positiever beoordeeld. Met name bij de slechte essays werd de schrijfster positiever beoordeeld als er een knappe foto bij zat. Bij goede essays is het effect van de foto veel kleiner. Dit is een voorbeeld van de interactie hypothese.

Kritiek op de interne validiteit van dit onderzoek is dat er niet is nagegaan of de manipulaties gelukt zijn. Er is wel een aselect toewijzing aan de condities (goed).

Wat betreft de externe validiteit zijn de proefpersonen voornamelijk mannelijke studenten (niet goed) en een foto is anders dan echte persoon (niet goed).

Er zijn repeated measures designs met meer dan twee condities:

- Equivalent time series: meerdere behandelingen worden getest op dezelfde units. Dit is een vorm van een quasi-experimenteel design. Voordelen zijn dat de standaardfouten kleiner zijn omdat je preciezer gaat meten. Nadelen zijn bijv. het conditie effect: stel, je laat mensen lichte oefeningen en zware oefeningen doen en meet hun depressieniveau na oefeningen. De lichte oefeningen hebben al een effect op depressie, maar je concludeert dat de zware oefeningen ook effect hebben omdat het depressieniveau daalt. Eigenlijk hebben alleen de lichte oefeningen effect gehad en voegden de zware oefeningen niets toe, je concludeert dan dus verkeerd.
- Latin square design: willekeurige volgorde van condities. Neem weer het voorbeeld van de oefeningen. Je hebt lichte, milde en zware oefeningen. De ene groep begint met lichte, dan zware en dan milde. De andere groep met milde, dan lichte, dan zware enz. Alle groepen krijgen dus alle condities, maar alleen in verschillende volgorden. De condities zijn dus ook random.

Interne validiteit:

Er zijn een aantal zaken die de interne validiteit kunnen bedreigen:

- Selection bias: je zet mensen in bepaalde groepen, dit maakt dat je geen random gekozen groepen hebt en dat kan ervoor zorgen dat de groepen niet representatief zijn.
- History: er gebeurt iets tijdens een experiment, dit kan invloed hebben op de resultaten in de nameting. Stel dat je depressie meet en de zon gaat schijnen tijdens het experiment, dan zal iemand waarschijnlijk minder hoog scoren op depressie in de nameting, omdat de zon iemand blij maakt.
- Maturation: binnen het experiment verandert er iets met bijv. een respondent waardoor de nameting anders is. Een respondent heeft bijvoorbeeld honger, hierdoor kan het zijn dat hij anders reageert.
- Testing: meerdere keren dezelfde test doen kan invloed hebben op de resultaten, doordat respondenten de vragenlijsten al kennen en dan anders gaan antwoorden.
- Experimental mortality: als mensen uit ene onderzoek wegvallen (mensen die afvallen gedurende therapieessie).

Bovenstaande problemen kunnen worden opgelost door de eerdergenoemde designs te gebruiken. Het beste design is dan een echt experiment.

Ook kan de interne validiteit bedreigd worden door:

- Experimenter of subjects expectancy zijn bepaalde verwachtingen. Bijvoorbeeld leerlingen met een hoog IQ zullen ook beter presteren. Dit komt uit. Verwachtingen hebben een effect op uitkomst (pigmalion effect).
Deze expectancy kan vermeden worden door een dubbelblinde opzet of een placebo. Dubbelblinde opzet betekent het contact via onderzoeksopzet waarin noch de experimenter noch de subjecten weten wie in welke groep zit. Placebo betekent dat de controlegroep een namaak-behandeling krijgt.
- Contamination of treatment betekent dat de controlegroep iets meekrijgt van de bedreiging. Wanneer de logica van de experimentele onderzoeksopzet goed is, kunnen we alternatieve verklaringen uitsluiten. Een mogelijke oplossing voor contamination of treatment is ervoor zorgen dat de groepen proefpersonen geen onderling contact hebben.

Externe validiteit:

Externe validiteit betekent dat de resultaten gegeneraliseerd moeten kunnen worden.

Bedreigingen:

- Een niet-representatieve steekproef zorgt ervoor dat het niet voor de hele populatie geldt. Daardoor kunnen de resultaten dan niet generaliseerd worden.
- Kunstmatigheid van de setting, kunstmatige manipulatie, reactiviteit van deelnemers. Een veldexperiment en een natuurlijk experiment kunnen hiervoor gebruikt worden als oplossing.