

Randomized Experiments

Rick Hoyle, Monica Harris, Charles Judd (2002), S. 237-269

In: Rick Hoyle, Monica Harris, Charles Judd (Hrsg.): *Research Methods in Social Relations*, 7. Auflage, ? u.a. (2002)

ABSTRACT

1) Das logische Designen von Experimenten soll der internen Validität dienen und da ist die randomisierte Zuordnung in Experimenten die strengste und beste Form. Für jedes Experiment benötigen wir abhängige und unabhängige Variablen. Abhängige werden beobachtet und unabhängige manipuliert – letztere auch im Feld durch externe Kontrolle. Unabhängige Variablen die nicht manipulierbar sind, aber dennoch unter den Individuen variieren werden als ‚individual differences‘ beschrieben. Im Idealexperiment werden zwei Gruppen gebildet (Treatment → Test) die einem mit und einmal ohne Veränderung der unabhängigen Variable den Test durchlaufen – bei mehreren Variablen könnte man nur eine variieren und die anderen festhalten, das senkt aber die interne Validität. Die Randomisierte Zuteilung erfolgt zu den Testgruppen und muss echt Zufällig (Gleichverteilt) sein. Die Ergebnisse verbessert der Forscher durch wiederholte Durchführung (repeatedly). Bei kleinen Gruppen können anstelle zwei Versuchsgruppen auch beide hintereinander mit und ohne der unabhängigen Variabel getestet werden (Varianz ‚within‘).

2) Interne Validität von Experimenten wird durch verschiedene Artefakte beeinflusst. Dazu zählen i) Selection [durch Auswahl der Gruppen entsteht Fehler → Lsg: durch Randomisierung der Zuteilung], ii) Maturation [Veränderung von Eigenschaften die auf die Variablen auswirken → Lsg: durch Kontrollgruppe], iii) History [Während dem Experiment geschieht extern etwas was die Ergebnisse intern beeinflusst → Lsg: multiple Gruppen für jede Variante], iv) Instrumentation [Das Verfahren hat Auswirkungen auf die Ergebnisse → Lsg: Training der Leiter oder Kontrollgruppe ohne Testing!], v) Mortality [Fortbleiben vom Test → Lsg: schwer, Motivieren], vi) Selection durch Maturation [Durch unterschiedliche Veränderung von Eigenschaften in der Gruppe werden ergebnisse beeinflusst → Lsg: Randomisierte Zuteilung]. Auch die Konstruktvalidität der operationalisierten Variablen ist problematisch, welche durch die kontinuierliche Veränderung & Kontrolle der Variablen verbessert werden kann. Sonst: Expertenmeinung.

3) Mehrere Versionen existieren, welche verschiedene Einflüsse beheben: i) Basic 2-Group design, ii) Randomized Two-Group Design, iii) Pretest-Posttest Two Group, iv) Solomon-four-way, v) Between-Participants Factorial Design, vi) Repeated Measures Design. Neben den Veränderungen im Design, welche die Validität erhöhen sollen und gegen die Artefakte der internen Validität gerichtet sind ist natürlich die gesamte Experimentalforschung dennoch Problematisch (obwohl klinische Studien so ablaufen). In den USA werden auch oft College-Studenten als Testpersonen genommen und dafür werden Experimentalkritiker auch kritisiert.

1) Design: Variablen & Zuteilung

INTRODUCTION

- “logical principles of design ... to increase the validity of the research” (S. 237)
- “research design of a study has a substantial effect on the study’s internal validity, that is, the ability to reach causal conclusions about the impact of one construct on another” (S. 237)
- “the strongest sort of designs from this point of view [internal validity], those are known as randomized experiments” (S. 237). Und diese dienen besonders “ideally suited for the task of causal analysis” (S. 237).

CONTROLLING VARIABLES

- “research requires the manipulation or measurement of variables” (S. 238) und “variables represent the construct that the researcher is interested in” (S. 238)
- “most research , we have ... at least two variables, and we look at the association between them” (S. 238). Und eben für diese Verbindungsmessung benötigen wir die “independent variable is the variable we believe to have a causal influence on our outcome variables” und natürlich die “dependent variable is the outcome variable” (S. 238). Wichtig ist, dass die unabhängigen Variablen “either they or someone else can manipulate” (S. 239). Diese werden auch präziser als „experimental variables“ (S. 239) bezeichnet.
- Es gibt noch eine andere Form der unabhängigen Variablen, und zwar solche, die man nicht manipulieren kann und die aber individuell sind: „individual differences variables. They are properties that people already possess“ (S. 239).
- Die Veränderungen im Design bei einer unabhängigen Variable sehen so aus, dass wir nun zwei Gruppen brauchen, wobei nur eine davon mit der unabhängigen Variable bestückt ist / bedient wird und die andere nicht. Die abhängige Variable soll dann in beiden Fällen beobachtet / erfragt werden.
- Im Falle mehrerer unabhängiger Variablen könnte man diese „holding them constant“ (S. 240). Aber das ist “never sufficient to be assumed of internal validity“ (S. 240), denn jede nicht geprüfte Variable als Unterscheidung zwischen Gruppen könnte als “plausible alternative explanation for our conclusion” (S. 240) dienen.

RANDOM ASSIGNMENT

- Es ist der einzige mögliche Weg um „to equate two or more groups on all possible individual difference variables at the start“ (S. 241). Und das Verfahren findet vor den eigentlichen Experimenten statt aber nach der Auswahl des Samples von der Population. Dafür kann das ‚Random Sampling‘ verwendet werden.
- Bei diesem Verfahren werden “way of assigning participants to the level of the independent variable so that the groups do not differ as the study begins“ (S. 241) und das natürlich auf eine Vielzahl von Experimenten die jeweils unterschiedliche Variablen überprüfen sollen. Ziel: “maximizes the internal validity of research” (S. 241).
- Die Zuweisung muss “truly random” (S. 241) sein und es darf nicht nach „basis of their arrival ... or alphabetically (S. 241) laufen. Z.B. bei zwei Gruppen für ein Sample „flipping with a coin“ (S. 242).
- Zur weiteren Erhöhung der internen Validität sollte das ‚random assignment‘ „repeatedly“ (S. 242) durchgeführt werden, also mehrmals hintereinander ein oder das Sample randomisiert den Experimentalgruppen zuweisen.
- Größte Gefahr sind hierbei die „failures of randomization“ (S. 243) also „even though randomization procedures were followed, the groups were not equivalent“ (S. 243).

INDEPENDENT VARIABLES VARY BETWEEN

- Bei kleinen Gruppen z.B. 30 Personen und einer unabhängigen Variable (z.B. Lärm) könnten wir nicht wollen, dass wir 15 Personen mit und 15 ohne Lärm testen, sondern wir würden gerne alle bei Lärm / nicht Lärm testen. Das geht als „repeated measures design“ (S. 244) = within-participants.
- Die ursprüngliche Form von Experimenten in z.B. zwei Gruppen einmal mit und einmal ohne die unabhängige Variable nennt man dann between participants
- „When the independent variable varies within participants, random assignment means that we randomly determine the order in which any given participant is exposed“ (S. 244) und anders herum “When the independent variable varies between participants, random assignment is accomplished by randomly assigning participants to one condition” (S. 244)

2) Validität: Interne & Konstruktbezogen

INTERNAL VALIDITY

- “alternative explanations are threats to the internal validity of the research proposition” (S. 245), es geht also um alternative Erklärungsmuster die bei randomisierten Experimenten auftreten können.
- Selection: “any preexisting difference between individuals in the different experimental conditions that can influence the dependent variable” (S. 245). Wenn als Beispiel Bungee-Jumping und Mental Health getestet wird, dann deselektieren sich alle Nicht-Bungeespringer „self-select into experimental conditions“ (S. 245). Lösung: „randomly assigned“ (S. 246)
- Maturation: „any naturally occurring process within persons that could cause a change in their behaviour“ (S. 246) also Wachstum, Langeweile, Motivation. Besonders bei Kleinkindern oder Studien über längere Zeiträume ist dieser Aspekt höchst problematisch. Lösung „easily cured by including a control group“ (S. 246).
- History: „any event that coincides with the independent variable and could affect the dependent variable” (S. 246) also “anything that happens during the course of the study that is unrelated to the independent variable” (S. 247). Lösung: “multiple groups for each condition” (S. 247) also z.B. 2x mit unabhängiger Variabel und 2x ohne.
- Instrumentation: “any change that occurs over time in measurement procedures or devices” (S. 247). Lösung: “careful training and monitoring of observers or measurement procedures” (S. 247).
- Mortality: “any attrition of participants from a study” (S.248) im Sinne des Fortbleibens, was natürlich bei “longitudinal research” (S. 248) problematisch ist. Außerdem ist auch die externe Validität gefährdet, weil die Begutachtung der Studien fällt. Lösung: „no easy cure“ (S. 248).
- Selection by Maturation: „when there are differences between individuals in the treatment groups that produces changes in the groups at different rates“ (S. 249) also unterschiedliche Veränderung von Aspekten die variablenbezogen sind. Lösung: “random assignment” (S. 249).
- Beispiel mit Auflistungen der einzelnen Probleme (S. 249-252).

CONSTRUCT VALIDITY

- “to control or manipulate the independent variable ... in laboratory settings ... can also be conducted in real-world settings” (S. 252).
- Nun muss aber diese Variabel die getestet werden soll auch messbar sein: “create an operational definition of it ... is the procedure used by the researcher to manipulate or measure the variables” (S. 252). Dies sind also “operational definitions of abstract concepts” (S. 252), wobei wichtig ist, dass “they are not unique to laboratory

experiments” (S. 252). Also auch für Feldexperimente werden operationalisierungen benötigt.

- Probleme die dabei auftauchen können sind “the independent variable is fairly complex, requiring a complex operational definition, or the independent variable is a general construct that can be operationally defined in any number of ways” (S. 253).
- ?? BEZUG ?? Als Maßzahl wie gut die Variable unser Konstrukt beschreiben, welches wir messen wollen wird ja die „construct validity“ (S. 253) verwendet, die „making sure that our variables capture the construct we wish to measure” (S. 253). “researcher must test the manipulation by subsequently measuring its effects to determine its construct validity” (S. 253), also durch ein Trail-n-Error-Verfahren Änderungen in der Operationalisierung austesten und sehen welche Ergebnisse sich einstellen bis die Richtung stimmt.

3) Alternativen & Schwächen

ALTERNATIVE EXPERIMENTAL DESIGNS

- Alle: Random Selection zur Ausschaltung von *Selection-Bias*
- Basic: “randomly assigned to the levels of the independent variable and each participant was measured only once” (S. 254) sowie “independent variable varies within participants and each participant is measured under each level” (S. 254).
- Randomized Two-Group Design: Zwei Gruppen mit randomisierter Zuordnung. Kontrollgruppe für Ausschaltung von *Instrumentation-Bias*. Treatment → Measure
- Pretest-Posttest Two Group: Zwei Gruppen mit randomisierter Zuordnung. Ausschaltung von *History-Bias*: Measure → Treatment → Measure
- Solomon-four-way: Vier Gruppen 2x mit Pretest → Treatment → Posttest und 2x Treatment → Posttest. Jeweils eine der beiden wird mit der unabhängigen Variable konfrontiert, so dass History, Interaktion, Testing ausgeschlossen werden können. Vereinfachung dieser Version bei Belasco & Tries 1969: 1x alles, 1x nur Nachtest (schafft auch alle Beeinflussungen auszugrenzen).
- Between-Participants Factorial Design: 2x2 = 4 Gruppen, 2x2x3 = ? Gruppen
- Repeated Measures Design: Nacheinander wiederholen

STRENGTH AND WEAKNESSES

- “strengths .. subsequent differences on the dependent variable are caused ... by the treatments and are not preexisting” (S. 264).
- Aber “no experiment can be 100% sure” (S. 264). Probleme werden hier “artefact” (S. 264) genannt und dazu zählen alle Probleme der internen Validität (vgl. 265), welche aber durch die Designs gefangen werden können (vgl. vorheriges Kapitel Solomon-4-way). Auch die „external validity“ (S. 265) ist ein Problemfaktor und das gilt für die gesamte Experimentalforschung: „extended beyond the laboratory“ (S. 265) ist äußerst fraglich. Ebenso gilt der Bias der ‚College Studenten‘ die in den USA oft für Samples herangezogen werden als hochproblematisch (vgl. S. 265 f.), weil sie nicht zufällig sind und auch nicht repräsentativ für die Population.