

2 Forschungsmethoden in der Sozialpsychologie

Andrew G. Livingstone und Antony S. R. Manstead

Zusammenfassung

Dieses Kapitel gibt einen Überblick über die Forschungsmethoden in der Sozialpsychologie, von der Entwicklung einer Theorie bis zur Datenerhebung. Nach der Beschreibung dreier quantitativer Forschungsstrategien (Experiment, Quasiexperiment und Umfrageforschung) wird kurz auf qualitative Ansätze eingegangen; der Schwerpunkt liegt dabei auf der Diskursanalyse. Danach werden die Schlüsselemente des Experiments behandelt, weil dies die am weitesten verbreitete Forschungsmethode in der Sozialpsychologie ist. Zudem beschäftigen wir uns mit Gefährdungen der Validität von Experimenten, der Wichtigkeit von Replikationen und den Problemen experimenteller Forschung in der Sozialpsychologie. Im letzten Abschnitt des Kapitels werden verschiedene Methoden der Datenerhebung dargestellt (Beobachtung, Selbstbeurteilung, implizite Messverfahren, Big-Data-Erfassung von Online-Verhalten und physiologische Messverfahren).

Learning Goals Start

Schlüsselbegriffe

- Abhängige Variable ✓
- Big Data ✗
- Cover Story ✓
- Diskursanalyse ✓
- Echtes Experiment mit Zufallszuweisung ✓
- Effekte von Versuchsleitungserwartungen ✓
- Einfache Zufallsstichprobe ✓
- Experiment ✓
- Experimentalgruppe ✗
- Experimentelle Konfundierung ✓
- Experimentelles Szenario ✓
- Externe Validität ✓
- Faktorielles Experiment ✓
- Feldexperiment ✓
- Haupteffekt ✓
- Hinweise aus der experimentellen Situation ✓
- Hypothese ✓
- Implizite Maße ✓
- Interaktionseffekt ✓
- Interne Validität ✓
- Konföderierte/r ✓
- Konstrukt ✓
- Konstruktvalidität ✓
- Kontrollgruppe ✗
- Konvergente Validität ✓
- Manipulationsüberprüfung ✓
- Mediatorvariable ✓
- Messwiederholungsdesign ✓
- Metaanalyse ✓
- Nichtreaktive Methoden ✓
- Nur-Nachtest-Kontrollgruppendesign ✓
- One-Shot-Fallstudie ✓
- Operationalisierung ✓
- Postexperimentelle Aufklärung ✓
- Postexperimentelle Befragung ✓
- Präregistrierung ✗

- Quasiexperiment ✓
- Quotenstichprobe ✓
- Reaktivität ✓
- Reliabilität ✓
- Replikation ✗
- Soziale Neurowissenschaft ✓
- Soziale Erwünschtheit ✓
- Stichprobenziehung ✓
- Teilnehmende Beobachtung ✓
- Theorie ✓
- Triangulation ✓
- Umfrageforschung ✓
- Unabhängige Variable ✓
- Validität ✓
- Variable ✓
- Versuchsperson ✗
- Zufallszuweisung ✓

Learning Goals Stop

Overview Start

Dieses Kapitel gibt einen Überblick über die Forschungsmethoden in der Sozialpsychologie, von der Entwicklung einer Theorie bis zur Datenerhebung. Nach der Beschreibung dreier quantitativer Forschungsstrategien (Experiment, Quasiexperiment und Umfrageforschung) wird kurz auf qualitative Ansätze eingegangen; der Schwerpunkt liegt dabei auf der Diskursanalyse. Danach werden die Schlüsselemente des Experiments behandelt, weil dies die am weitesten verbreitete Forschungsmethode in der Sozialpsychologie ist. Zudem beschäftigen wir uns mit Gefährdungen der Validität von Experimenten, der Wichtigkeit von Replikationen und den Problemen experimenteller Forschung in der Sozialpsychologie. Im letzten Abschnitt des Kapitels werden verschiedene Methoden der Datenerhebung dargestellt (Beobachtung, Selbstbeurteilung, implizite Messverfahren, Big-Data-Erfassung von Online-Verhalten und physiologische Messverfahren).

Overview Stop

2.1 Einleitung

Questions Start

Wie werden in der Sozialpsychologie Theorien überprüft?

Questions Stop

Warum werden Forschungsmethoden in diesem Lehrbuch bereits im zweiten Kapitel thematisiert? Warum sollten wir uns erst mit Forschungsmethoden beschäftigen, anstatt mit den sozialpsychologischen Phänomenen selbst und deren Erklärungen zu beginnen? Um diese Fragen zu beantworten, müssen wir eine sogar noch grundlegendere Frage beantworten: Warum machen wir überhaupt Forschung?

In der Sozialpsychologie interessieren wir uns natürlich für die „großen“ Phänomene: Wie entsteht ein Intergruppenkonflikt? Warum belegen Menschen Mitglieder anderer Gruppen mit Stereotypen? Wie bilden wir uns Eindrücke von anderen Menschen? Warum verhalten sich Leute anders, wenn sie in einer Gruppe sind? Was bringt Menschen dazu, ihre Einstellungen zu ändern? Von welchen Faktoren hängt es ab, ob enge Beziehungen gelingen oder scheitern? Um Fragen wie diese zu beantworten, entwickeln wir *Theorien*. Beispielsweise wollen wir vielleicht eine **Theorie** über die Ursachen für Intergruppenkonflikte entwickeln. Zunächst einmal gehört dazu, die relevanten **Konstrukte** zu benennen, also abstrakte Begriffe wie etwa „Bedrohung“ oder „Vorurteil“, von denen angenommen

wird, dass sie für die Fragestellung von Belang sind. Dann müssen dazugehörige **Variablen** definiert werden, also messbare Repräsentationen dieser Konstrukte wie Testwerte bei Fragebogen zu Bedrohungswahrnehmungen oder zur Intergruppenfeindseligkeit. Schließlich treffen wir Annahmen darüber, wie die betreffenden Konstrukte miteinander zusammenhängen. Im Kern bestehen unsere Theorien typischerweise aus Annahmen über *Kausalzusammenhänge* zwischen Konstrukten. So geben wir uns nicht damit zufrieden, diese „großen“ Phänomene einfach nur zu *beschreiben*. Vielmehr versuchen wir, sie dadurch zu *erklären*, dass wir ihre Anfangsbedingungen identifizieren. Bei der Entwicklung einer Theorie über Intergruppenkonflikte (► Kap. 14) interessieren wir uns nicht nur dafür, worin ein Konflikt besteht, sondern auch dafür, was ihn verursacht und wie er abgebaut werden könnte. Beispielsweise könnten wir die folgende Theorie aufstellen: Ein Intergruppenkonflikt wird durch den Eindruck verursacht, dass die Interessen oder das Wohlergehen der eigenen Gruppe durch eine andere Gruppe bedroht sind (s. Branscombe, Ellemers, Spears & Doosje, 1999).

Definition Start

Definition

Theorie (theory): Ein System aus abstrakten Begriffen (d. h. Konstrukten) und Aussagen darüber, wie diese Konstrukte miteinander zusammenhängen.

Definition Stop

Definition Start

Definition

Konstrukt (construct): Ein abstrakter theoretischer Begriff (z. B. „sozialer Einfluss“).

Definition Stop

Definition Start

Definition

Variable (variable): Messbare Repräsentation eines Konstrukts.

Definition Stop

Als Basis unserer Theorien könnten wir entweder Beobachtungen aus dem realen Leben oder unsere Intuition verwenden oder wir könnten auf bereits bestehende Theorien zurückgreifen. Theorien aufzustellen ist jedoch nur eine Seite der Medaille. In vielen anderen Disziplinen wie Philosophie, Soziologie und Anthropologie beschäftigen sich Forschende mit Themen und Phänomenen, die auch in der Sozialpsychologie von Interesse sind. Zur Abgrenzung der Sozialpsychologie – aber auch der Psychologie insgesamt – von diesen anderen Disziplinen trägt nicht nur die *Art* der Erklärung bei, die wir liefern, sondern auch die Festlegung auf die *wissenschaftliche Methode*, mit der wir unsere Theorien anhand von *Befunden* überprüfen. Dies bedingt eine zentrale Eigenschaft einer Theorie: Sie muss *überprüfbar* sein. Das bedeutet, dass wir aus der Theorie spezifische Vorhersagen (d. h. **Hypothesen**) über die Zusammenhänge zwischen zwei oder mehreren Konstrukten ableiten können, die sich in empirischen Untersuchungen auch als falsch herausstellen können. Wird eine Theorie durch die Befundlage gestützt, sollte sie beibehalten werden. Sollten die Befunde der Theorie jedoch widersprechen, müssen wir die Theorie hinterfragen oder sogar zurückweisen. Eine bestehende Theorie sollte immer dann aktualisiert oder durch eine neue Theorie ersetzt werden, wenn auf letztere zumindest eine der folgenden Eigenschaften zutrifft: (a) Die neue Theorie kann die verfügbaren Daten besser oder genauer erklären; (b) sie liefert eine solche Erklärung mit weniger Annahmen oder Kernkonstrukten (d. h., sie ist sparsamer); (c) sie erklärt sowohl bereits vorhandene als auch *neue* Befunde, für die die alte Theorie noch keine Erklärung hatte; oder (d) sie regt weiterführende Forschung an, indem sie uns zu besseren Forschungsfragen und neuen Hypothesen

verhilft (für eine allgemeine Diskussion über Theorien und deren Evaluation, s. Tzeng & Jackson, 1991; Meehl, 1990, bietet ebenfalls eine klassische, aber recht anspruchsvolle Diskussion darüber, wie wir Theorien evaluieren können).

Definition Start

Definition

Hypothese (hypothesis): Eine aus einer Theorie abgeleitete Vorhersage über die Zusammenhänge zwischen Variablen.

Definition Stop

Nehmen wir als Beispiel die Theorie von Janis (1982) über die schlechte Qualität von Entscheidungsprozessen, wie sie manchmal sogar in Gruppen mit kompetenten und erfahrenen Personen zu beobachten ist; hier handelt es sich um ein Phänomen, das als „**Gruppendenken**“ (groupthink) bezeichnet wird (► Kap. 8). Die Theorie von Janis (1982) umfasst (a) Anfangsbedingungen schlechten Entscheidungsverhaltens in Gruppen, (b) Symptome des Gruppendenkens, (c) Symptome schlechten Entscheidungsverhaltens und schließlich (d) Annahmen über den Prozess, der die Anfangsbedingungen mit den Symptomen des Gruppendenkens und des schlechten Entscheidungsverhaltens verknüpft (* Abb. 2.1). Eine der *Anfangsbedingungen* ist eine „kohäsive Gruppe“, also eine Gruppe, deren Mitglieder psychologisch von der Gruppe abhängig sind. Aufgrund dieser Abhängigkeit ist es wahrscheinlich, dass ihre Mitglieder mit dem, was sie als Konsens in der Gruppe ansehen, übereinstimmen. Ein Beispiel für ein *Symptom* von Gruppendenken ist die Anwesenheit von „Gedankenwachen“ (mind guards). Mit diesem Ausdruck bezeichnet Janis Gruppenmitglieder, die es sich zur Aufgabe machen, die Gruppe vor Informationen abzuschirmen, die die sachliche Richtigkeit bzw. moralische Legitimität einer sich herauskristallisierenden Entscheidung infrage stellen würden. Ein Beispielsymptom fehlerhaften Entscheidungsverhaltens ist das Versagen, die Risiken der bevorzugten Entscheidung abzuwägen. Janis spezifizierte auch im Detail, wie Gruppendenken zustande kommt (er beschrieb also den *vermittelnden Prozess*). Demnach besteht der vermittelnde Prozess in einer voreiligen „concurrency-seeking tendency“, also einem starken Streben nach Übereinstimmung mit anderen Gruppenmitgliedern, bevor alle für die Entscheidung wichtigen Fragen wirklich gründlich erörtert worden sind. Somit verknüpft die Theorie die Anfangsbedingungen von Gruppendenken über einen vermittelnden Prozess mit Symptomen; wir werden uns weiter unten genauer mit dem Begriff „vermittelnder Prozess“ (Mediation, Mediatorvariablen) beschäftigen.

Platzhalter Abbildung Start

Abb. 2.1 Anfangsbedingungen, vermittelnder Prozess und Symptome des Gruppendenkens im theoretischen Modell von Janis (1982, © Free Press/Simon & Schuster, © Cengage Learning)

Datei:

Bildrechte: [Urheberrecht beim Autor]

Abdruckrechte: Nicht notwendig

Hinweise Verlag/Setzerei: alt: 24802_6_De_2_Fig1_Print.eps, **Korrektur: Abb. 2.1** Beschriftung im Kreis in der Mitte: „TENDENZ ZUR ÜBEREINSTIMMUNG MIT ANDEREN GRUPPENMITGLIEDERN“ wie im Text. Box auf der rechten Seite: „Gedankenwachen“ statt der männlichen „-wächter“.

Platzhalter Abbildung Stop

Logisch lässt sich die folgende Vorhersage aus der Theorie von Janis ableiten: Gruppen, die sich durch ein hohes Maß an **Kohäsion** (► Kap. 12) auszeichnen, treffen eher schlechte Entscheidungen als Gruppen mit einer geringeren Kohäsion (► Aus der Forschung: Auswertung von Archivdaten zum „Gruppendenken“). Je stärker die Befunde mit einer Vorhersage übereinstimmen, desto mehr Vertrauen können wir in die zugrunde liegende Theorie haben, aus der die Vorhersage abgeleitet wurde. Entsprechend sollten wir weniger Vertrauen in die zugrunde liegende Theorie haben, wenn die Befunde nicht mit der Vorhersage übereinstimmen. Befunde können auch Grenzen oder

einschränkende Bedingungen für einen vorhergesagten Effekt aufzeigen und nahelegen, dass dieser nur unter bestimmten Bedingungen auftritt. Ein Beispiel für einschränkende Bedingungen bezogen auf Janis' Theorie geht auf die Forschung von Postmes, Spears und Cihangir (2001) zurück. Sie fanden heraus, dass es von der *Norm* der Gruppe abhängt, wie stark der Effekt der Gruppenkohäsion auf die Qualität der Gruppenentscheidung ist: Die Entscheidungsqualität verbesserte sich, wenn die Gruppe die Norm kritischen Denkens befolgte; sie verschlechterte sich, wenn die Norm, den Konsens aufrechtzuerhalten, im Vordergrund stand. Angesichts solcher Befunde muss eine ursprüngliche Theorie möglicherweise modifiziert oder sogar zugunsten einer Alternativtheorie zurückgewiesen werden.

Jetzt verstehen wir, warum Forschungsmethoden so wichtig sind: Sie sind die wesentlichen Instrumente, um unsere Ideen in Handeln umzusetzen (z. B. um ein Experiment zu planen) und um unsere Theorien zu überprüfen, kritisch zu hinterfragen und zu verbessern. Die Qualität unserer Forschung beruht nicht nur auf der Qualität unserer Theorien, sondern auch auf der Qualität der Forschungsmethoden, mit denen wir diese Theorien überprüfen.

Results Start

Zusammenfassung

Methoden sind die Werkzeuge, die Forschende verwenden, um ihre theoretischen Vorstellungen zu überprüfen. Diese Vorstellungen können aus einer ganzen Reihe von Quellen stammen, zwei davon sind jedoch in der Sozialpsychologie weitverbreitet: Beobachtungen aus dem realen Leben und Widersprüche zwischen Befunden vorangegangener Forschung. Eine Theorie besteht aus systematisch miteinander verbundenen Konstrukten und trifft Vorhersagen über das Auftreten bestimmter Phänomene.

Results Stop

Case Study Start

Aus der Forschung

Auswertung von Archivdaten zu „Gruppendenken“

Janis, I. L. (1972). *Victims of groupthink: A psychological study of foreign-policy decisions and fiascoes*. Boston: Houghton Mifflin.

Einleitung

Janis' Forschung zum Thema Gruppendenken ist ein hervorragendes Beispiel für die Auswertung von Archivdaten, eine Forschungsstrategie, die ansonsten im vorliegenden Kapitel nicht beschrieben oder erörtert wird. Bei der Archivforschung stammen die Daten aus Archiven, d. h. aus gespeicherten Aufzeichnungen über Fakten. „Zu Archivdaten können Materialien gehören wie persönliche Dokumente (Briefe oder Tagebücher), kreative Werke (Gedichte, Gemälde, Aufsätze), Biografien oder Autobiografien sowie historische Dokumente oder Aufzeichnungen der Regierung“ (Simonton, 1981, S. 218). Janis (1972) entschied sich, Archivmaterialien im Zusammenhang mit vier bedeutenden Fehlschlägen der US-amerikanischen Außenpolitik im Detail zu untersuchen, und zwar Materialien über (1) die Invasion in der kubanischen Schweinebucht 1961, (2) die Entscheidung, den Koreakrieg im Jahre 1950 auszuweiten, (3) die Unterschätzung des Gegners, die 1941 zum japanischen Überraschungsangriff auf Pearl Harbor führte, und (4) die Entscheidung, 1964 den Vietnam-Krieg auszuweiten. Nach Janis' Ansicht standen den verantwortlichen Politikern bei jeder einzelnen dieser katastrophalen Entscheidungen potenziell oder tatsächlich Informationen zur Verfügung, die zu anderen Entscheidungen hätten führen müssen.

Methode

Janis ging bei seiner Untersuchung so vor, dass er alle dokumentarischen Informationsquellen über die Umstände durchforstete, unter denen diese fehlerhaften Entscheidungen gefällt wurden. In seinem

Buch *Victims of Groupthink* aus dem Jahre 1972 versuchte er zu zeigen, wie die Archivdaten über jede einzelne dieser Entscheidungen im Sinne eines konsistenten sozialpsychologischen Musters interpretiert werden können, dessen Kern in Abb. 2.1 dargestellt ist. In einer überarbeiteten Auflage seines Buchs (Janis, 1982) wendete er seine Auffassung von Gruppendenken auf den Watergate-Skandal an. Der Watergate-Skandal führte 1974 schließlich zum Rücktritt des damaligen amerikanischen Präsidenten Richard Nixon.

Spätere Forschung

Tetlock (1979) führte eine stärker quantitativ ausgerichtete Auswertung von Archivmaterial durch. Er nutzte standardisierte Verfahren zur Auswertung des Inhalts öffentlicher Äußerungen von Schlüsselpersonen, die an Entscheidungen mit und ohne Gruppendenken, wie sie Janis (1972) untersucht hatte, beteiligt gewesen waren. Tetlock wollte vor allem erfassen, wie stark die öffentlichen Äußerungen dieser Politiker die Tendenz zum Ausdruck brachten, „politikrelevante Informationen auf vereinfachte und verzerrte Weise zu verarbeiten“ (S. 1317) sowie „die eigene Gruppe als sehr positiv und Personen aus dem gegnerischen Lager als sehr negativ zu bewerten“ (S. 1317).

Um diese beiden Aspekte des Gruppendenkens zu erfassen, identifizierte er sechs Schlüsselpersonen, die an fünf unterschiedlichen außenpolitischen Entscheidungen direkt beteiligt gewesen waren; zwei dieser Entscheidungen hatte Janis als „Entscheidungen ohne Gruppendenken“ klassifiziert, die anderen drei hingegen als „Entscheidungen mit Gruppendenken“. Dann wählte Tetlock aus öffentlichen Erklärungen, die von diesen Schlüsselpersonen zur Zeit der betreffenden Krise abgegeben worden waren, jeweils zufällig zwölf Textpassagen in der Länge eines Absatzes aus und analysierte sie. Er fand heraus, dass die öffentlichen Äußerungen der Schlüsselpersonen in Krisen mit Gruppendenken signifikant weniger komplex waren als diejenigen in Krisen ohne Gruppendenken. Er fand auch Belege dafür, dass in Krisen mit Gruppendenken die Entscheidungsverantwortlichen positivere Bewertungen ihrer eigenen politischen Gruppen abgaben, als dies in Krisen der Fall war, die nicht durch Gruppendenken gekennzeichnet waren. Im Widerspruch zu den Vorhersagen äußerten sich die Schlüsselpersonen in Krisen ohne Gruppendenken ähnlich negativ gegenüber Personen aus dem gegnerischen Lager wie die Schlüsselpersonen in Krisen mit Gruppendenken. Mit Ausnahme dieses letzten Befunds stehen die Ergebnisse der Studie von Tetlock im Einklang mit den Schlussfolgerungen von Janis, die auf einer eher qualitativen Analyse historischer Dokumente beruhen.

Allerdings ergaben andere Archivstudien weniger Unterstützung für Janis' Theorie. Peterson, Owens, Tetlock, Fan und Martorana (1998) nutzten Archivmethoden, um die Entscheidungsfindung im oberen Management von sieben führenden US-Unternehmen (z. B. Coca-Cola) zu untersuchen. Sie interessierten sich dafür, ob die Symptome des Gruppendenkens bei erfolglosem Management häufiger auftraten als bei erfolgreichem Management. Haslam (2004, S. 102) fasst die Ergebnisse so zusammen: „Im Gegensatz zu Janis' Hypothese tauchten die Symptome des Gruppendenkens anscheinend in *allen* Gruppenentscheidungen auf und waren nicht charakteristischer für erfolgloses Management als für erfolgreiches.“

Diskussion

Der Hauptvorteil der Strategie der Archivforschung besteht darin, dass das aus den Archiven zusammengetragene Datenmaterial nicht durch das Wissen der Versuchspersonen darüber, dass Forschende deren Verhalten untersuchen, verzerrt ist. Das in diesem archivierten Material gezeigte Verhalten erfolgte in einem natürlichen Kontext zu einem früheren Zeitpunkt, d. h. vor Beginn der Untersuchung des Materials. Es besteht deshalb nur eine geringe oder gar keine Gefahr, dass das Verhalten durch den Forschungsprozess „kontaminiert“ sein könnte. Simonton (1981) drückte es so aus: „Weil die Archivforschung Daten auswertet, die schon von anderen für Zwecke gesammelt wurden, die sich oft deutlich von den Absichten der Forschenden unterscheiden, handelt es sich hierbei um eine nichtreaktive Forschungsmethode“ (S. 218). Diesem Vorteil stehen jedoch auch einige Nachteile gegenüber. Die offensichtlichsten sind die folgenden: (1) Die Forschenden sind abhängig von der Qualität der Archivdaten, die eventuell keine gute Grundlage für die Erfassung der zentralen Variablen darstellen; (2) es ist nicht klar, ob und wie die gemessenen Variablen (beispielsweise die Entscheidungsqualität und die Komplexität der Aussagen der an der Entscheidung Beteiligten) *kausal*

zusammenhängen.

Case Study Stop

2.2 Forschungsstrategien

Questions Start

Was sind die Stärken und die Schwächen der wichtigsten Forschungsstrategien, die in der Sozialpsychologie zur Verfügung stehen?

Questions Stop

Zur Überprüfung ihrer Vermutungen und Vorhersagen stehen Forschenden eine Vielfalt unterschiedlicher Forschungsstrategien zur Verfügung. In diesem Abschnitt gehen wir näher auf die experimentelle und quasiexperimentelle Forschung, auf die Umfrageforschung sowie auf qualitative Ansätze ein.

2.2.1 Experimente und Quasiexperimente

Die experimentelle Forschung dient dem Nachweis kausaler Beziehungen. Das Ziel eines **Experiments** besteht darin, zu beobachten, was mit einem Phänomen geschieht, wenn Forschende gezielt ein bestimmtes Merkmal der Umwelt verändern, in der das Phänomen auftritt („Werden sich Veränderungen bei Variable *A* ergeben, wenn ich Variable *B* verändere?“). Indem sie die Veränderungen der Variable *B* kontrollieren, können Forschende aus den beobachteten Veränderungen der Variable *A* kausale Schlussfolgerungen über den Zusammenhang zwischen *A* und *B* ziehen. Statt einfach nur zu wissen, dass ein Mehr von Variable *A* mit einem Mehr von Variable *B* einhergeht, entdecken experimentell Forschende, ob *A* zunimmt, wenn *B* zunimmt, *A* abnimmt, wenn *B* verringert wird, *A* stabil bleibt, wenn *B* unverändert gelassen wird usw. Ein solches Ergebnismuster würde darauf hindeuten, dass Veränderungen von *B* Veränderungen von *A* *verursachen*.

Definition Start

Definition

Experiment (experiment): Methode, bei der die Versuchsleitung gezielt eine Veränderung einer Situation herbeiführt, um die Folgen dieser Veränderung zu untersuchen.

Definition Stop

Die experimentelle Methode kann vielfältig variieren. Zwei verbreitete Variationen sind das **Quasiexperiment** und das **echte Experiment mit Zufallszuweisung**. Sie unterscheiden sich voneinander in der Echtheit der Situation, in der die Daten erhoben werden, und im Ausmaß an Kontrolle, die die Forschenden über die Situation ausüben. Ein Quasiexperiment wird normalerweise in einer natürlichen, alltäglichen Umgebung durchgeführt, über die die Forschenden keine vollständige Kontrolle besitzen. Das echte Experiment mit Zufallszuweisung zeichnet sich hingegen dadurch aus, dass die Versuchsleitung die Schlüsselmerkmale der Situation vollständig kontrolliert; dies bringt jedoch oft eine gewisse Künstlichkeit mit sich.

Definition Start

Definition

Quasiexperiment (quasi-experiment): Ein Experiment, bei dem die Versuchspersonen den verschiedenen Experimentalbedingungen nicht zufällig zugeteilt werden (sondern in der Regel aufgrund von Faktoren, die nicht der Kontrolle der Forschenden unterliegen).

Definition Stop

Definition Start

Definition

Echtes Experiment mit Zufallszuweisung (true randomized experiment): Ein Experiment, bei dem die Versuchspersonen den unterschiedlichen Experimentalbedingungen zufällig zugewiesen werden.

Definition Stop

Um den grundlegenden Unterschied zwischen einem Quasiexperiment und einem echten Experiment besser nachvollziehen zu können, müssen wir uns noch intensiver mit dem Begriff „*Experiment*“ auseinandersetzen. Experimente sind Untersuchungen, in denen Forschende die Auswirkungen einer Klasse von Variablen (unabhängige oder manipulierte Variablen) auf eine andere Klasse von Variablen (abhängige oder gemessene Variablen) untersuchen. In einem echten Experiment mit Zufallszuweisung haben Forschende *sowohl* Kontrolle über die unabhängige Variable *als auch* darüber, wer mit dieser Variable konfrontiert wird. Besonders wichtig ist es dabei, dass die Forschenden in der Lage sind, die **Versuchspersonen** zufällig auf die unterschiedlichen Bedingungen des Experiments zu verteilen (**Zufallszuweisung**). In einem Quasiexperiment können Forschende meistens keinen Einfluss darauf nehmen, mit welchen experimentellen Bedingungen die jeweiligen Versuchspersonen konfrontiert werden.

Definition Start

Definition

Versuchsperson (participant): Person, die an einer psychologischen Studie teilnimmt.

Definition Stop

Definition Start

Definition

Zufallszuweisung (random allocation, random assignment): Der Vorgang, bei dem eine Versuchsperson einer Versuchsbedingung so zugeordnet wird, dass sie mit der gleichen Wahrscheinlichkeit auch jeder der anderen Versuchsbedingungen zugeordnet werden könnte.

Definition Stop

Die Unterschiede zwischen den beiden Methoden lassen sich an Beispielen verdeutlichen. In der Aggressionsforschung ist die Frage von besonderem Interesse, ob Filme mit Gewaltdarstellungen einen Einfluss auf das spätere Verhalten der Zuschauenden haben (► Kap. 9). Diese Frage kann sowohl im Rahmen von echten Experimenten mit Zufallszuweisung als auch mittels Quasiexperimenten untersucht werden. Ein Beispiel für ein echtes Experiment zu diesem Thema ist die Studie von Liebert und Baron (1972): Jungen und Mädchen in zwei Altersgruppen wurden zufällig einer von zwei Experimentalbedingungen zugewiesen. Während in der einen Bedingung Filmausschnitte mit Gewaltszenen zu sehen waren, wurden in der anderen Bedingung Ausschnitte aus einem spannenden Sportwettkampf gezeigt. Danach hatten beide Gruppen von Kindern vermeintlich die Möglichkeit, einem anderen Kind weh zu tun. Die Kinder, die die Gewaltszenen gesehen hatten, nutzten diese Möglichkeit häufiger als die Kinder, denen das gewaltfreie Material gezeigt worden war. Da die Kinder den beiden Experimentalbedingungen nach dem Zufallsprinzip zugewiesen worden waren, können die beobachteten Unterschiede auf das unterschiedliche Filmmaterial zurückgeführt werden statt auf vorher bestehende Unterschiede zwischen den Kindern.

Ein Beispiel für eine quasiexperimentelle Untersuchung zum selben Thema ist die Studie von Black und Bevan (1992): Hier füllten Personen unter einer von vier Bedingungen einen Fragebogen aus, der

die Bereitschaft zu aggressivem Verhalten messen sollte: (1) während sie vor dem Kino in einer Schlange anstanden, um einen Gewaltfilm zu sehen, (2) während sie anstanden, um einen gewaltfreien Film zu sehen, (3) nachdem sie gerade einen Gewaltfilm gesehen hatten oder (4) nachdem sie gerade einen gewaltfreien Film gesehen hatten. Wie Sie • Abb. 2.2 entnehmen können, fanden Black und Bevan (1992) heraus, dass diejenigen, die auf einen Gewaltfilm warteten, eine größere Neigung zu aggressivem Verhalten zeigten als diejenigen, die auf einen gewaltfreien Film warteten. Es stellte sich ferner heraus, dass Befragte, die gerade einen Gewaltfilm gesehen hatten, eine größere Tendenz zur Aggression zeigten als diejenigen, die warteten, um einen Gewaltfilm zu sehen. Es fanden sich jedoch keine Unterschiede zwischen denen, die gerade einen gewaltfreien Film gesehen hatten, und denen, die auf einen gewaltfreien Film warteten. Diese Befunde stehen im Einklang mit der Vermutung, dass Gewaltfilme die Aggressionsbereitschaft erhöhen (• Abb. 2.3). Durch den Umstand, dass die Versuchspersonen den unterschiedlichen Bedingungen nicht zufällig zugewiesen worden waren, lassen sich allerdings auch andere Erklärungsmöglichkeiten nicht ausschließen. Beispielsweise könnte es sein, dass Gewaltfilme die Aggressionsbereitschaft nur bei denjenigen Personen erhöhen, die überhaupt an derartigen Filmen interessiert sind.

Platzhalter Abbildung Start

Abb. 2.2 Die selbst beurteilte Neigung, sich aggressiv zu verhalten, in Abhängigkeit von der Art des Films und davon, ob die befragten Personen darauf warteten, einen Film zu sehen, oder ihn gerade gesehen hatten (Daten von Black & Bevan, 1992)

Datei:

Bildrechte: [Urheberrecht beim Autor]

Abdruckrechte: Nicht notwendig

Hinweise Verlag/Setzerei: alt: 24802_6_De_2_Fig3_Print.eps

Platzhalter Abbildung Stop

Platzhalter Abbildung Start

Abb. 2.3 Nimmt bei uns allen die Neigung zu aggressivem Verhalten zu, wenn wir einen gewalthaltigen Film gesehen haben, oder nehmen aggressive Neigungen von vornherein nur bei denjenigen zu, die sich zum Ansehen solcher Filme hingezogen fühlen? (© imago / United Archives)

Datei:

Bildrechte: [Urheberrecht beim Autor]

Abdruckrechte: Nicht notwendig

Hinweise Verlag/Setzerei: alt: 24802_6_De_2_Fig4_Print.eps

Platzhalter Abbildung Stop

Oftmals ist ein Quasiexperiment die einzige Möglichkeit, eine experimentelle Studie zu einem sozialen Phänomen durchzuführen. Ethische und praktische Erwägungen machen es häufig unmöglich, Personen verschiedenen experimentellen Bedingungen *zufällig* zuzuweisen. Wenn man sich etwa wie Stroebe, Stroebe und Schut (2007) für die Wirkungen eines Trauerfalls auf die Gesundheit interessiert, ist es offensichtlich nicht möglich, die untersuchten Personen zufällig den beiden Bedingungen „verwitwet“ und „nicht verwitwet“ zuzuweisen. Dieses Problem zeigt sich auch in vielen anderen Forschungsgebieten. Die Wahl der Forschungsstrategie ist somit oft ein Kompromiss zwischen dem Optimalen und dem praktisch Möglichen. Erfreulicherweise ist die Entwicklung mancher quasiexperimenteller Designs so weit fortgeschritten, dass es mit einiger Gewissheit möglich ist, aus ihnen kausale Schlussfolgerungen zu ziehen (Judd & Kenny, 1981a, b; West, Biesanz & Pitts, 2000).

Es ist auch im Feld möglich, echte Experimente durchzuführen. Solche **Feldexperimente** (► Kap. 1) kombinieren die Vorteile beider Ansätze: die Kontrollierbarkeit eines Laborexperiments und den Realismus eines Quasiexperiments. Ein Beispiel für ein solches Feldexperiment findet sich in ► Aus der Forschung: Wie sich mithilfe der Medien Vorurteile abbauen und Konflikte zwischen Gruppen reduzieren lassen: Ein Feldexperiment in Ruanda.

Definition Start

Definition

Feldexperiment (field experiment): Ein echtes Experiment mit Zufallszuweisung, das in einem natürlichen Setting durchgeführt wird.

Definition Stop

Case Study Start

Aus der Forschung

Wie sich mithilfe der Medien Vorurteile abbauen und Konflikte zwischen Gruppen reduzieren lassen:
Ein Feldexperiment in Ruanda

Paluck, E. L. (2009). Reducing intergroup prejudice and conflict using the media: A field experiment in Rwanda. *Journal of Personality and Social Psychology*, 96, 574–587.

Einleitung

Das Ziel dieser Studie von Paluck war es, die Wirksamkeit einer Intervention zur Verbesserung der Beziehungen zwischen sozialen Gruppen in Ruanda zu untersuchen (* Abb. 2.4). 1994 wurden in Ruanda im Zuge eines Völkermords etwa 800.000 Angehörige der Tutsi-Minderheit von der Hutu-Mehrheit getötet. Die von Paluck untersuchte Intervention erfolgte in Form einer für das Radio produzierten Seifenoper. Diese enthielt Botschaften zum Abbau von Vorurteilen zwischen Tutsi und Hutu und sollte den Zuhörenden zudem die Symptome eines psychischen Traumas besser verständlich machen. Die beiden zentralen Forschungsfragen waren dabei, (1) ob das Radioprogramm tatsächlich einen Einfluss auf Vorurteile und Gruppenkonflikte nehmen konnte, und falls ja, (2) wie dieser Einfluss aussehen mochte. Paluck war insbesondere daran interessiert, ob die Radiointervention die folgenden drei Variablen beeinflussen konnte: (1) die vorurteilsbehafteten *Überzeugungen* von Individuen, (2) die wahrgenommenen *Normen* innerhalb ihrer Gemeinden (d. h., was normal und akzeptabel ist) und/oder (3) beobachtbares Verhalten innerhalb der Gemeinden in der Form der Beteiligung an Diskussionen über Intergruppenbeziehungen und Trauma. Zur Überprüfung wurden die Zuhörenden der Seifenoper mit einer Kontrollgruppe verglichen, die stattdessen ein Radioprogramm über Gesundheitsfragen hörte.

Methode

Die Versuchspersonen kamen aus zwölf verschiedenen Gemeinden, die so ausgewählt wurden, dass sie eine repräsentative Stichprobe der ruandischen Gesellschaft darstellten: Acht der Gruppen setzten sich aus der Allgemeinbevölkerung aus vier unterschiedlichen Regionen zusammen, zwei bestanden aus Überlebenden des Völkermords und die letzten beiden waren Twa (eine weitere Minderheit in Ruanda). Aus jeder Kategorie sollte die Hälfte der Gemeinden jeweils das zur Versöhnung produzierte Radioprogramm hören, während der anderen Hälfte das Kontrollprogramm (mit dem Fokus auf Gesundheit) vorgespielt wurde. Während eines Jahres besuchte das Forschungsteam einmal pro Monat jede Gruppe und spielte die neue Episode des jeweiligen Programms ab. Die Versuchspersonen versammelten sich und hörten die Episode gemeinsam in der Gruppe an. Nach einem Jahr besuchte Paluck mit ihrem Team jede der Gruppen während drei Tagen, um die Ergebnisse der Intervention an einer Zufallsstichprobe von 40 Versuchspersonen zu untersuchen. Dabei verwendeten sie verschiedene Methoden wie etwa Einzelinterviews, Fokusgruppen und Verhaltensbeobachtungen. Die Einzelinterviews beinhalteten Fragen zur Einschätzung vorurteilsbehafteter Überzeugungen der Individuen, zu den wahrgenommenen Normen innerhalb ihrer Gemeinde in Bezug auf die Interaktion mit anderen Gruppen und zur Empathie für andere Gruppen, einschließlich ruandischer Häftlinge (darunter auch Täter des Genozids) und Personen, die den Genozid überlebt hatten. Bei der Verhaltensbeobachtung wurden Interaktionsmuster kodiert, die sich während des Radiohörens in der Gruppe zeigten. Gezählt wurden die Äußerungen abweichender Meinungen (d. h., wie oft sich die Personen an einer Diskussion beteiligten, statt sich an der von einer

Autoritätsperson geäußerten Meinung zu orientieren) sowie die positiven Äußerungen über Gruppenkooperation.

Ergebnisse

Im Vergleich zum Kontrollprogramm wirkte sich die Seifenoper, die auf Versöhnung ausgerichtet war, positiv auf die Zuhörenden aus. Die Wirkung der Seifenoper war jedoch beschränkt auf die wahrgenommenen Normen innerhalb der Gemeinde, wohingegen für individuelle vorurteilsbehaftete Überzeugungen kein Effekt festgestellt werden konnte. So stimmten die Versuchspersonen in der Experimentalgruppe stärker solchen Aussagen zu wie „Mischehen sollten in meiner Familie erlaubt sein“ und „Es ist nicht naiv, anderen zu vertrauen“. Beide Aussagen beziehen sich auf Normen dafür, wie sich die Mitglieder einer Gemeinde verhalten sollten. Im Gegensatz dazu hatte die Intervention keinen Einfluss auf die Zustimmung zu Aussagen wie „Wer in einer Gewaltsituation untätig zusieht, trägt Verantwortung“ und „Mischehen führen zu Frieden“ – Aussagen also, die die Überzeugungen der Individuen hinsichtlich der Beziehung zwischen den Gruppen widerspiegeln. Personen, die das versöhnliche Radioprogramm gehört hatten, zeigten darüber hinaus mehr Empathie für andere ruandische Gruppen, äußerten häufiger abweichende Meinungen (was auf die Bereitschaft hindeutet, traditionelle Autoritätspersonen infrage zu stellen) und gaben positivere Kommentare zur Gruppenkooperation ab.

Diskussion

Insgesamt deuten diese Ergebnisse darauf hin, dass Medieninterventionen wie das hier verwendete Radioprogramm zur Versöhnung das Potenzial haben, die Beziehungen zwischen Gruppen – selbst vor dem traumatischen Hintergrund eines Völkermords – zu verbessern. Die Ergebnisse geben zudem Aufschluss darüber, wie dies geschieht: Die Intervention wirkte in erster Linie auf wahrgenommene Gemeinenormen über typisches und wünschenswertes Verhalten und nicht auf persönliche Vorurteile. Die Studie von Paluck bietet somit wichtige theoretische Erkenntnisse darüber, wie Beziehungen zwischen Gruppen verbessert werden können. Konkret deutet sie darauf hin, dass die Vermittlung sozialer Normen über die Beziehungen zu anderen Gruppen erfolgsversprechender zu sein scheint als der Versuch, die Ansichten von Individuen über andere Gruppen zu ändern. Außerdem ist die Studie ein Paradebeispiel einer Feldstudie. Durch das Setting des gemeinsamen Radiohörens ist sie sehr realitätsnah. Gleichzeitig verwendet sie strenge Methoden: Die Stichprobenziehung erfolgte nach Gesichtspunkten der Repräsentativität und bei der Einteilung in die experimentellen Bedingungen kam eine Form der Zufallszuweisung zum Einsatz. Schließlich veranschaulicht die Studie das Potenzial der Sozialpsychologie, positive gesellschaftliche Auswirkungen zu haben – und dies selbst unter den verheerendsten sozialen Bedingungen.

Platzhalter Abbildung Start

Abb. 2.4 Kann eine Seifenoper dazu beitragen, dass sich die Menschen in Ruanda nach dem Völkermord wieder versöhnen?

Datei:

Bildrechte: [Urheberrecht beim Autor]

Abdruckrechte: Nicht notwendig

Hinweise Verlag/Setzerei: neues Foto

Platzhalter Abbildung Stop

Case Study Stop

2.2.2 Umfrageforschung

Die praktischen und ethischen Überlegungen, die für Quasiexperimente sprechen, lassen sich in einem allgemeineren Sinne auch auf die **Umfrageforschung** übertragen (Oppenheim, 1992; Schwarz, Groves & Schuman, 1998; Visser, Krosnick & Lavrakas, 2000). Umfragen unterscheiden sich insofern von Experimenten und Quasiexperimenten, als sie darauf abzielen, das bestehende Niveau relevanter Variablen zu erfassen, anstatt sie zu manipulieren (* Abb. 2.5). Wie bei Experimenten geht es in

Umfragen meist darum, Zusammenhänge und/oder Ursache-Wirkungs-Beziehungen zwischen Variablen zu finden. Aufgrund der fehlenden Kontrolle über die unabhängigen Variablen – sie werden auch *Prädiktoren* genannt – ist es jedoch praktisch nicht möglich, Gewissheit über ihre kausale Rolle zu erlangen. Aus diesem Grund werden in Umfragen oft andere Variablen erfasst, die in die statistischen Auswertungen einbezogen werden können – die dann also *statistisch kontrolliert* werden –, um mögliche Alternativerklärungen für den Zusammenhang zwischen einer Prädiktor- und einer Ergebnisvariable auszuschließen.

Platzhalter Abbildung Start

Abb. 2.5 Eine Strategie zum Sammeln von Forschungsbefunden besteht darin, die öffentliche Meinung mithilfe von Interviews zu erfassen. (© Josef Muellek / Thinkstock)

Datei:

Bildrechte: [Urheberrecht beim Autor]

Abdruckrechte: Nicht notwendig

Hinweise Verlag/Setzerei: alt: 24802_6_De_2_Fig9_Print.eps; neue Abb.

Platzhalter Abbildung Stop

Definition Start

Definition

Umfrageforschung (survey research): Eine Forschungsstrategie, bei der eine Stichprobe von Befragten interviewt wird (oder einen Fragebogen ausfüllt); die Befragten werden so ausgewählt, dass sie repräsentativ für die Population sind, aus der sie gezogen wurden.

Definition Stop

Ein Beispiel für diese Strategie ist die Studie von Pratto, Sidanius, Stallworth und Malle (1994). Sie vertraten die Auffassung, dass viele unterschiedliche Arten von Intergruppenvorurteilen durch eine Persönlichkeitsvariable vorhergesagt werden können, die sie als *soziale Dominanzorientierung* bezeichneten (social dominance orientation, kurz: SDO; ► Kap. 14). Diese Variable bringt zum Ausdruck, inwiefern Menschen eine Vorliebe für ungleichen Status zwischen sozialen Kategorien haben: Eine Person, die hohe Werte in Bezug auf SDO hat, sollte beispielsweise sexistischer sein und stärker gegen gleiche Rechte für ethnische Minderheiten eingestellt sein. Im Einklang mit dieser Vorhersage fanden Pratto et al. (1994), dass SDO in der Tat positiv mit Sexismus korrelierte und negativ mit der Unterstützung der Gleichberechtigung aller ethnischen Gruppen. Die Forschenden beschäftigten sich jedoch auch mit der Frage, ob der Zusammenhang zwischen diesen Variablen und SDO möglicherweise auf dem Einfluss einer weiteren Variable beruhte, nämlich politischem Konservatismus. Sollte dies der Fall sein, wäre das Argument entkräftet, dass SDO einen (wichtigen) Prädiktor für Vorurteile darstellt. Um diese Möglichkeit zu überprüfen, werteten Pratto et al. erneut die Korrelationen zwischen Vorurteilen und SDO aus, indem sie den Effekt des politischen Konservatismus statistisch kontrollierten (d. h. *auspartialisierten*). Die Korrelationen blieben signifikant; dies spricht dafür, dass der gefundene Zusammenhang nicht einfach nur auf dem politischen Konservatismus beruht.

Definition Start

Definition

Stichprobenziehung (sampling): Auswahl einer Teilmenge von Personen aus einer Population mit der Absicht, die Population zu beschreiben, aus der sie gezogen wurde.

Definition Stop

Bei der Umfrageforschung werden zwei Arten der **Stichprobenziehung** verwendet, probabilistische und nichtprobabilistische. Die grundlegendste probabilistische Form der Stichprobenziehung ist die

einfache Zufallsstichprobe. Bei einer einfachen Zufallsstichprobe müssen zwei Bedingungen erfüllt sein: Erstens hat jedes einzelne Mitglied der Population die gleiche Chance, ausgewählt zu werden; zweitens hat jede mögliche Kombination der gewünschten Anzahl Mitglieder die gleiche Ziehungswahrscheinlichkeit. Um die zweite Bedingung zu erklären, stellen wir uns vor, dass die Größe der Population 10 (sie besteht aus Personen A bis J) und die Größe der Stichprobe 2 ist. Es gibt 45 mögliche Kombinationen von zwei Mitgliedern der Population (A + B, A + C, A + D usw. bis I + J). Bei einer einfachen Zufallsstichprobe muss jede einzelne dieser 45 möglichen Kombinationen von zwei Mitgliedern gleich wahrscheinlich sein. In der Praxis ist eine Zufallsstichprobe (z. B. der gesamten Bevölkerung eines Landes) viel größer und der Prozess viel komplexer. Forschende ziehen eine Zufallsstichprobe, indem sie jedem einzelnen Mitglied der Population eine Zahl zuweisen und computergenerierte Zufallszahlen verwenden, um eine Stichprobe der erforderlichen Größe zu ziehen; hierbei handelt es sich um ein Verfahren, das im Internet zur Verfügung steht.

Definition Start

Definition

Einfache Zufallsstichprobe (simple random sample): Eine Stichprobe, die so gezogen wird, dass jedes Mitglied der Population die gleiche Chance hat, ausgewählt zu werden, und jede mögliche Kombination der gewünschten Anzahl von Mitgliedern die gleiche Ziehungswahrscheinlichkeit hat.

Definition Stop

Weil die probabilistische Stichprobenziehung teuer und zeitaufwendig ist, wird häufig eine nichtprobabilistische Stichprobenziehung angewendet. Die am weitesten verbreitete Form der nichtprobabilistischen Stichprobe ist die **Quotenstichprobe**. Hier besteht das Ziel darin, eine Stichprobe auszuwählen, in der die grundlegenden Merkmale der Population repräsentiert werden. Solche Merkmale können Alter und Geschlecht sein. Wenn die Alters- und die Geschlechterzusammensetzung der betreffenden Population bekannt ist, ist es möglich, die Alters- und die Geschlechterzusammensetzung der Stichprobe der Population anzupassen. Der Begriff „Quote“ bezieht sich auf die Anzahl der Menschen mit bestimmten Merkmalen (z. B. Frauen im Alter zwischen 55 und 60 Jahren), die interviewt werden müssen. Der Hauptvorteil der Quotenstichprobe besteht darin, dass bei der Rekrutierung alle Personen angefragt werden können, die die erwähnten Quotenmerkmale aufweisen, ohne dass eine bestimmte, vorher identifizierte Person rekrutiert werden muss. Nachteile von Quotenstichproben sind, (1) dass die Interviewenden entscheiden müssen, an wen sie sich wenden, um eine Quote zu erfüllen, wodurch Verzerrungen bei der Auswahl auftreten können, und (2) dass es unmöglich ist, eine genaue Schätzung des Stichprobenfehlers anzugeben.

Definition Start

Definition

Quotenstichprobe (quota sample): Eine Stichprobe, die bestimmten, im Vorhinein festgelegten Quoten entspricht und damit bestimmte Merkmale der Population (wie Alter oder Geschlecht) widerspiegelt, die als für die Forschungsfrage relevant angesehen werden.

Definition Stop

Eine Strategie, mit der in der Umfrageforschung das Vertrauen in die Annahme eines Kausalzusammenhangs gestärkt werden kann, besteht darin, die relevanten Variablen zu mehreren Zeitpunkten zu erfassen – in einem *längsschnittlichen* Umfragedesign. Dahinter steckt folgende Logik: Wenn sich aus Variable A zum Zeitpunkt 1 Variable B zum Zeitpunkt 2 vorhersagen lässt (vor allem wenn die Unterschiede in Bezug auf Variable B zum Zeitpunkt 1 statistisch kontrolliert werden), können wir mit größerer Wahrscheinlichkeit annehmen, dass Variable A eine kausale Wirkung auf Variable B hat.

2.2.3 Qualitative Ansätze

Üblicherweise wird in der sozialpsychologischen Forschung eine *quantitative* Datenanalyse durchgeführt; das gilt auch für die große Mehrheit der Studien, die in diesem Buch vorgestellt werden. Damit ist gemeint, dass die ausgewerteten Daten durch Zahlen repräsentiert werden. Auf diese Weise werden unsere Daten für statistische Auswertungen zugänglich und es wird den Forschenden möglich, etwas über die für eine Stichprobe typische Ausprägung einer Variable auszusagen (z. B. das arithmetische Mittel, den Median und/oder den Modalwert), über die *Streuung* der Werte einer Variable (z. B. die Varianz oder Standardabweichung) und die *Stärke* und *Zuverlässigkeit* der Zusammenhänge zwischen zwei oder mehr Variablen (z. B. mithilfe inferenzstatistischer Verfahren wie *t*-Test, Varianzanalyse oder Regression). Diese Auswertungen liefern meistens klare und interpretierbare Ergebnisse; zudem besteht unter den Forschenden ein breiter Konsens über die Regeln bei der Interpretation von Ergebnissen.

Der quantitative Ansatz kann jedoch in der Sozialpsychologie bei Weitem keine universelle Verbreitung beanspruchen. Eine Alternative ist die *qualitative* Analyse, bei der die Daten typischerweise eher aus Texten als aus Zahlen bestehen; analysiert werden dabei Inhalt und Bedeutung der von den untersuchten Personen verwendeten Sprache bzw. Wörter. Qualitative Ansätze umfassen eine breite Vielfalt analytischer Techniken. Einige davon stehen im Einklang mit den philosophischen Annahmen, die der quantitativen Forschung bzw. der wissenschaftlichen Methode im Allgemeinen zugrunde liegen. Zu diesen Annahmen gehört die Überzeugung, dass die Phänomene, für die wir uns interessieren, eine objektive Menge an „Tatsachen“ darstellen. Nach dieser Auffassung existieren diese Phänomene unabhängig von der Sichtweise, die die Forschenden (oder irgendeine andere Person) auf sie haben. Demnach brauchen wir zur Untersuchung von Phänomenen, die nicht direkt beobachtet oder gemessen werden können – wie etwa eine soziale Einstellung – in erster Linie geeignete und durchdachte Forschungsmethoden. Andere qualitative Ansätze sind allerdings radikaler und lehnen diese Annahmen explizit ab. Was qualitative Ansätze jedoch eint, ist die Auffassung, dass die quantitative Erforschung sozialer Phänomene an Grenzen stößt und dass qualitative Techniken zusätzliche oder sogar völlig andere Einsichten liefern können (s. Henwood, 1996).

Quantitativen Methoden wird vorgeworfen, sie würden Phänomene und die Sicht der untersuchten Personen auf diese Phänomene möglicherweise falsch repräsentieren oder unangemessen vereinfachen. Insbesondere verlangen quantitative Maße – vor allem *Selbstbeurteilungsmaße*, auf die wir später noch kommen werden – von Forschenden, dass sie Annahmen über den Umfang und den Inhalt möglicher Reaktionen machen; dabei sprechen wir noch gar nicht von der *Bedeutung*, die die Reaktionen für die an einer Studie teilnehmenden Personen haben. Solche Annahmen haben Konsequenzen: Die Versuchspersonen verstehen unter Umständen eine Aufgabe oder ein Fragebogenitem anders als die Forschenden. Darüber hinaus schränken die von den Forschenden vorgesehenen Antwortalternativen die Teilnehmenden möglicherweise ein und verunmöglichen ihnen, bestimmte Reaktionen zu zeigen oder ihre Sichtweisen auszudrücken.

Verschiedene qualitative Ansätze tragen zu einer Lösung dieses Problems bei. In frühen, explorativen Phasen eines Forschungsprojekts können die *Inhaltsanalyse* oder die *thematische Analyse* freier mündlicher oder schriftlicher Antworten zu neuen Erkenntnissen über potenziell relevante Faktoren führen. Dieser Ansatz lässt sich auch als Grundlage für die Exploration unerwarteter oder mehrdeutiger Befunde nutzen (zu einem Beispiel hierfür s. Livingstone & Haslam, 2008). Tatsächlich führen Inhaltsanalysen am Ende oft zu *quantitativen* Ergebnissen. Denn sie ermöglichen es, zu zählen, wie häufig bestimmte Wörter, Sätze oder Themen vorkommen; damit lassen sich dann statistische Auswertungen durchführen. Im Unterschied dazu versucht die *grounded theory*, induktiv oder „datengeleitet“ eine Theorie über ein bestimmtes Phänomen zu *konstruieren*, beispielsweise auf der Grundlage explorativer Interviewdaten. Andere Techniken (z. B. die *interpretative phänomenologische Analyse*) konzentrieren sich darauf, die subjektive Bedeutung zu eruieren und zu interpretieren, die aus dem Blickwinkel der Versuchspersonen bestimmten Themen oder Ereignissen anhaftet.

Alle diese Methoden können die quantitative Forschung ergänzen oder erweitern. Im Gegensatz dazu werden in anderen qualitativen Ansätzen wie manchen Varianten der **Diskursanalyse** grundsätzlich

die Annahmen abgelehnt, die der quantitativen Forschung (und einigen qualitativen Ansätzen) zugrunde liegen. Anstatt nach objektiv erkennbaren „Tatsachen“ über sozialpsychologische Phänomene zu suchen, nehmen die Vertreterinnen und Vertreter dieser Ansätze an, dass es nicht nur eine gültige Interpretation der Welt gibt. Konsequenterweise legen sie den Fokus ihrer Forschung auf Merkmale der Alltagssprache und alltäglicher Interaktion, um zu erkunden, wie Menschen aktiv bestimmte Interpretationen von Ereignissen konstruieren. Sie versuchen beispielsweise zu zeigen, dass rassistische oder sexistische Einstellungen nicht auf individuellen Überzeugungen oder Verzerrungen beruhen, sondern auf Bewertungen, die im Kontext bestimmter sozialer Interaktionen entstehen. Demnach handelt es sich bei derartigen Bewertungen nicht um die unveränderlichen Produkte individueller kognitiver Systeme; vielmehr entstehen sie im Kontext von Gesprächen und hängen von der jeweiligen sozialen Situation ab.

Definition Start

Definition

Diskursanalyse (discourse analysis): Eine Familie von Methoden, mit denen sich Gespräche und Texte analysieren lassen. Das Ziel dabei besteht darin, herauszufinden, wie Menschen ihrer Alltagswelt Sinn zuschreiben.

Definition Stop

Ein Beispiel für die Verwendung der Diskursanalyse findet sich bei Wetherell, Stiven und Potter (1987). Diese Forschenden interessierten sich für die Ansichten Studierender zu den beruflichen Möglichkeiten von Frauen. Sie erwarteten, dass in solchen Ansichten die Ideologien zum Vorschein kommen würden, die im Alltag Ungleichheiten zwischen den Geschlechtern reproduzieren. Mit 17 Studierenden wurde ein halbstrukturiertes Interview durchgeführt; ihre Antworten wurden transkribiert und analysiert. Ein Vorteil dieses Vorgehens besteht darin, dass es die Forschenden in die Lage versetzte, Widersprüche in der Art und Weise festzustellen, wie ganz normale Menschen über Themen wie die Ungleichheit zwischen den Geschlechtern sprechen. Anstatt eine einzige Einstellung zu vertreten, neigten die Studierenden dazu, zu unterschiedlichen Zeitpunkten während des Interviews unterschiedliche Positionen einzunehmen. Einige dieser Positionen waren untereinander inkonsistent, dienten aber zu dem jeweiligen Zeitpunkt, zu dem sie eingenommen wurden, speziellen ideologischen und strategischen Zwecken.

Die Diskursanalyse und verwandte Ansätze wie die Konversationsanalyse werden im vorliegenden Buch nicht behandelt. Dennoch stellt die Diskursanalyse eine wichtige und eindrucksvolle Forschungstradition dar, die soziale Phänomene typischerweise auf eine andere Art und Weise angeht und zu anderen Erkenntnissen führt als der Großteil der in diesem Buch besprochenen Forschung. Akademische Zeitschriften wie *Research on Language and Social Interaction* berichten über qualitativ hochwertige Diskurs- und Konversationsanalyseforschung. Zudem bietet eine 2012 veröffentlichte Sonderausgabe des *British Journal of Social Psychology* einen guten Einblick in die Art der Forschung, die mit der Diskursanalyse betrieben wird, und die philosophischen Annahmen, auf denen der Ansatz im Allgemeinen beruht. Es ist dabei wichtig zu betonen, dass sich die Annahmen der Diskursanalyse auch sehr deutlich von denen der meisten anderen qualitativen Forschungsmethoden unterscheiden, die ein sehr breites Spektrum an unterschiedlichen Techniken abdecken und die die in diesem Kapitel skizzierten quantitativen Ansätze mehr oder weniger ergänzen können.

Wie unsere Übersicht verschiedener Forschungsstrategien zeigt, besteht eine der wichtigsten Entscheidungen, die wir während des Forschungsprozesses fällen, in der Festlegung auf eine der oben skizzierten Strategien. Sicher sind einige Forschungsstrategien zur Untersuchung eines bestimmten Phänomens besser geeignet als andere; allerdings hat jede Strategie, wie gelungen ihre Umsetzung auch immer sein mag, ihre Grenzen. Aus diesem Grund hat sich einer der großen Pioniere der Forschungsmethodologie in den Sozialwissenschaften, Donald T. Campbell, für die **Triangulation** ausgesprochen. Damit meinte er, dass mehrere unterschiedliche Methoden bei der Untersuchung einer bestimmten Frage eine bessere Grundlage für eindeutige Schlussfolgerungen

darstellen als eine dieser Methoden alleine. Weil jede Methode ihre eigenen Stärken und Schwächen hat, trägt die Verwendung unterschiedlicher Methoden dazu bei, dass die Stärken der einen Methode die Schwächen der anderen ausgleichen (z. B. Fine & Elsbach, 2000).

Definition Start

Definition

Triangulation (triangulation): Die Verwendung *mehrerer* Methoden und Messinstrumente, um ein gegebenes Thema zu erforschen.

Definition Stop

Results Start

Zusammenfassung

Forschungsstrategien sind umfassende Kategorien von Forschungsmethoden, die uns zur Verfügung stehen, um sozialpsychologische Phänomene zu untersuchen. Zu Beginn haben wir angemerkt, dass es oft sinnvoll ist, ein Phänomen mithilfe von mehr als einer Strategie zu untersuchen. Bevor wir uns mit qualitativen Forschungsstrategien beschäftigten, skizzierten wir drei quantitative Strategien: das Experiment, das Quasiexperiment und die Umfrageforschung.

Results Stop

2.3 Näheres zum Experiment in der Sozialpsychologie

Questions Start

Was sind die Hauptbestandteile eines sozialpsychologischen Experiments?

Questions Stop

Wie zuvor erwähnt, ist das Experiment die vorherrschende Forschungsmethode der Sozialpsychologie. Dies liegt vor allem daran, dass diese Methode hinsichtlich der Überprüfung von Kausalannahmen über Beziehungen zwischen Variablen konkurrenzlos ist. Standardlehrbücher der sozialpsychologischen Forschung (z. B. Dunn, 2012) behandeln das Experiment meist als bevorzugte Forschungsmethode. Wie wir noch sehen werden, sind jedoch tatsächlich einige begründete Zweifel daran angebracht, ob sich aus experimentellen Studien eindeutige kausale Schlussfolgerungen ableiten lassen.

Wir werden zunächst die Hauptmerkmale des experimentellen Ansatzes in der Sozialpsychologie beschreiben. Dazu werden wir als anschauliches Beispiel Milgrams (1965; ► Kap. 8) bekannte Studie zum Gehorsam heranziehen.

2.3.1 Merkmale des sozialpsychologischen Experiments

Das **experimentelle Szenario** ist der Kontext, in dem den Versuchspersonen die Studie präsentiert wird. In einer Laborsituation ist es wichtig, ein Szenario zu entwickeln, für das es eine überzeugende und gut durchdachte Begründung gibt. Denn die Situation sollte den Versuchspersonen realistisch erscheinen und sie dazu bringen, sich wirklich engagiert zu beteiligen. Zudem sollten die experimentellen Manipulationen und der Messvorgang den Versuchspersonen nicht „ins Auge springen“. Im Milgram-Experiment wurde den Versuchspersonen gesagt, dass es sich bei der Studie um eine Untersuchung zu den Auswirkungen von Bestrafung auf das Lernen handele. Der Versuchsperson wurde, scheinbar zufällig, die Rolle des „Lehrers“ zugewiesen, während ein Komplize des Versuchsleiters (ein sogenannter **Konföderierter**) als weitere Versuchsperson auftrat und die Rolle des „Schülers“ übernahm. Die Aufgabe des Schülers bestand darin, eine Liste von Wortpaaren auswendig zu lernen. Aufgabe des Lehrers war es, jeweils das erste Wort eines Paares laut vorzulesen

und dann zu prüfen, ob sich der Schüler richtig an das zugehörige zweite Wort erinnerte. Wenn ihm dies nicht gelang (der Schüler war dazu instruiert worden, von Zeit zu Zeit Fehler zu machen), sollte ihn der Lehrer mit einer Folge elektrischer Stromstöße von ansteigender Spannung bestrafen. Das Szenario sollte die Versuchsperson davon überzeugen, dass die Stromstöße echt waren und der Schüler tatsächlich die Stromstöße erhielt (was in Wahrheit nicht zutraf).

Definition Start

Definition

Experimentelles Szenario (experimental scenario): „Verpackung“, in der ein Experiment den Versuchspersonen dargeboten wird.

Definition Stop

Definition Start

Definition

Konföderierte/r (confederate): Hierbei handelt es sich scheinbar um eine weitere Versuchsperson, die in Wirklichkeit jedoch mit der Versuchsleitung „verbündet“ ist und eine genau vorgeschriebene Rolle im Experiment spielt.

Definition Stop

Die **unabhängige Variable (UV)** ist die Variable, die die Versuchsleitung gezielt manipuliert. Alle anderen Aspekte des experimentellen Szenarios werden konstant gehalten, und die unabhängige Variable wird systematisch variiert.

Definition Start

Definition

Unabhängige Variable (independent variable): Variable, die in einer Studie systematisch variiert (manipuliert) wird, um ihren Einfluss auf eine oder mehrere abhängige Variablen zu untersuchen.

Definition Stop

Operationalisierung bezeichnet die Art und Weise, wie eine Variable konkret manipuliert bzw. gemessen wird. Im Milgram-Experiment war eine der wichtigsten unabhängigen Variablen die räumliche Nähe zwischen Lehrer und Schüler. In einer der Bedingungen befanden sie sich in getrennten Räumen. In einer zweiten Bedingung konnte der Lehrer den Schüler hören, aber nicht sehen. In der dritten Bedingung konnte er die Reaktionen des Schülers sowohl sehen als auch hören. In der vierten Bedingung musste der Lehrer die Hand des Schülers auf eine Metallplatte drücken, damit der Stromstoß erfolgen konnte. Alle anderen Aspekte der experimentellen Situation wurden konstant gehalten, sodass alle Unterschiede im Verhalten des Lehrers zwischen diesen vier Bedingungen ausschließlich auf Veränderungen der Nähe zwischen Lehrer und Schüler zurückführbar waren.

Definition Start

Definition

Operationalisierung (operationalization): Die Art und Weise, mit der in einer bestimmten Studie ein theoretisches ▶ Konstrukt in eine messbare ▶ abhängige Variable oder eine manipulierbare ▶ unabhängige Variable überführt wird.

Definition Stop

Die Qualität eines Experiments hängt häufig entscheidend davon ab, wie wirksam die Manipulationen der unabhängigen Variable sind. Unter *Wirksamkeit* verstehen wir

- das Ausmaß, in dem die Veränderungen der unabhängigen Variable die wesentlichen Eigenschaften des Konstrukts erfassen, von dem eine Theorie annimmt, dass es einen kausalen Einfluss auf das Verhalten hat, und
- die Stärke der eingeführten Veränderungen.

Betrachten wir beispielsweise, wie adäquat die vier Nähebedingungen in Milgrams Experiment das Konstrukt der Nähe repräsentieren: Was ganz offensichtlich manipuliert wurde, ist die *physische* Nähe. Zudem geht es um die Frage, ob die Veränderungen zwischen den vier Bedingungen hinreichend groß waren, um einen Effekt hervorzurufen. In diesem Fall ist es schwer vorstellbar, auf welche Weise die Variable Nähe drastischer hätte manipuliert werden können. Bei schwächeren Manipulationen besteht die Gefahr, die vorhergesagten Effekte nicht nachweisen zu können, einfach weil die Unterschiede zwischen den Stufen der unabhängigen Variable zu gering sind, um einen Effekt zu haben. In sozialpsychologischen Experimenten ist es mittlerweile gängige Praxis, zusätzlich zur Erfassung der abhängigen Variablen Messungen für die Wirksamkeit der Manipulation zu erheben; diese Messungen werden als **Manipulationsüberprüfungen** bezeichnet.

Definition Start

Definition

Manipulationsüberprüfung (manipulation check): Ein Maß dafür, wie erfolgreich die unabhängige Variable implementiert wurde.

Definition Stop

Definition Start

Definition

Abhängige Variable (dependent variable): Variable, von der erwartet wird, dass sie sich in Abhängigkeit von den Veränderungen der unabhängigen Variable verändert. Gemessene Veränderungen der abhängigen Variable (AV) werden als „abhängig“ von den manipulierten Veränderungen der unabhängigen Variable (UV) angesehen.

Definition Stop

Um die Auswirkungen einer unabhängigen Variable beurteilen zu können, ist es notwendig, Aspekte des Verhaltens oder des inneren Zustands der Versuchspersonen zu erfassen. Diese gemessene Variable wird als **abhängige Variable** (AV) bezeichnet, weil systematische Veränderungen dieser Variable von der Wirkung der unabhängigen Variable *abhängig* sein sollten. Im Milgram-Experiment war die abhängige Variable die Intensität der vom Lehrer verabreichten Stromstöße; zur Verfügung standen dem Lehrer 30 Stufen. Die Ergebnisse des Milgram-Experiments werden häufig als Prozentsatz der Versuchspersonen dargestellt, die die maximale Schockhöhe wählten, nämlich 450 Volt. Die Ergebnisse von Milgrams Studie (1965) sind in • Abb. 2.6 dargestellt. Die Schlüsselfrage im Hinblick auf jede abhängige Variable lautet, ob sie eine gute Messung des zugrunde liegenden theoretischen Konstrukts darstellt. Die Frage der Passung zwischen einem theoretischen Konstrukt und der gemessenen abhängigen Variable wird in ► Abschn. 2.3.3 noch detaillierter erörtert.

Platzhalter Abbildung Start

Abb. 2.6 Prozentsatz der Versuchspersonen, die die maximale Stromstoßstärke anwendeten und die deshalb als völlig gehorsam klassifiziert wurden. (Nach Milgram, 1965)

Datei:

Bildrechte: [Urheberrecht beim Autor]

Abdruckrechte: Nicht notwendig

Hinweise Verlag/Setzerei: alt: 24802_6_De_2_Fig8_Print.eps; **Korrektur:** Bitte auch in der Abbildung „teilnehmer“ in „personen“ ändern!

Platzhalter Abbildung Stop

Zu Laborexperimenten gehört oft ein gewisses Maß an Täuschung; damit ist gemeint, dass die Versuchspersonen über einige Aspekte der Untersuchung in die Irre geführt werden. Das Ausmaß dieser Täuschung kann vom einfachen Zurückhalten von Informationen über den wahren Zweck der Untersuchung bis zur absichtlichen Irreführung der Versuchspersonen reichen. Unter Umständen macht die Versuchsleitung diese also glauben, dass die Untersuchung ganz andere Ziele verfolge, als es in Wirklichkeit der Fall ist. Der hauptsächliche Grund für den Einsatz von Täuschung besteht darin, dass die Versuchspersonen anders handeln würden, wenn sie vom wahren Ziel der Untersuchung wüssten. Wenn Milgrams Versuchspersonen gewusst hätten, dass es sich um eine Studie zum Gehorsam handelte, wäre es mit Sicherheit häufiger zu Ungehorsam gekommen: Die Versuchspersonen hätten vermutlich demonstrieren wollen, dass sie sich dem Befehl widersetzen können, einen Mitmenschen zu verletzen.

Seit Milgrams Untersuchung haben sich die Einstellungen zur Verwendung von Täuschung in der sozialpsychologischen Forschung gewandelt: Die Irreführung von Versuchspersonen über den eigentlichen Zweck eines Experiments wird heute negativer gesehen. Dieser Wandel hat teils ethische Gründe (wo immer möglich, sollten wir es vermeiden, andere Personen zu täuschen, sei es im Kontext eines Experiments oder außerhalb) und teils praktische Gründe (wenn Versuchspersonen ständig in die Irre geführt werden, werden sie bei jeder künftigen Studie von der Erwartung ausgehen, dass sie erneut in die Irre geführt werden; das verändert vermutlich ihr Verhalten). Es ist schwierig, beiden Zielen gerecht zu werden: einerseits gegenüber den Versuchspersonen völlig ehrlich zu sein, andererseits sie zu untersuchen, ohne dass ihr Verhalten durch das Wissen über Art und Zweck eines Experiments beeinflusst wird. Die psychologische Forschung, die an den Universitäten in Europa, Nordamerika und Australasien durchgeführt wird, unterliegt typischerweise der Bewilligung einer „Ethikkommission“. Die Ethikkommission nimmt eine Einschätzung und eine Kontrolle von Forschungsaktivitäten vor, an denen Menschen als Versuchspersonen beteiligt sind. Darüber hinaus haben beispielsweise die American Psychological Association (APA), die British Psychological Association und die Deutsche Gesellschaft für Psychologie (DGPs) Richtlinien herausgegeben, an die sich Forschende halten sollten, wenn sie Untersuchungen mit Menschen als Versuchspersonen durchführen. Auf den Internetseiten der nationalen psychologischen Berufsorganisationen finden Sie die in Ihrem Land gültigen vollständigen ethischen Richtlinien; in • Tab. 2.1 präsentieren wir aus den ethischen Richtlinien der DGPs die Abschnitte zu Forschung und Publikation.

Tab. 2.1 Grundsätze des Berufsverbandes Deutscher Psychologinnen und Psychologen und der Deutschen Gesellschaft für Psychologie für Forschung und Publikation (Aus: DGPs, 2016, S. 22 ff., im Internet: https://www.dgps.de/fileadmin/user_upload/PDF/berufsethik-foederation-2016.pdf)

7.3 Grundsätze für Forschung und Publikation

(1)	<i>Forschung mit Menschen</i> Psychologische Forschung ist auf die Teilnahme von Menschen als Versuchspersonen angewiesen. Psychologinnen und Psychologen sind sich der Besonderheit der Rollenbeziehung zwischen Versuchsleiterin bzw. Versuchsleiter und Versuchsteilnehmerin bzw. Versuchsteilnehmer und der daraus resultierenden Verantwortung bewusst. Sie stellen sicher, dass durch die Forschung Würde und Integrität der teilnehmenden Personen nicht beeinträchtigt werden. Sie treffen alle geeigneten Maßnahmen, Sicherheit und Wohl der an der Forschung teilnehmenden Personen zu gewährleisten, und versuchen, Risiken auszuschließen.
(2)	<i>Förmliche Bewilligungen</i> Falls Forschungsprojekte einer förmlichen ethischen Bewilligung unterliegen, liefern Psychologinnen und Psychologen präzise Informationen über ihr Forschungsvorhaben. Sie

Kommentiert [ST1]: an Verlag: Abdruckgenehmigung erforderlich!

Deutsche Gesellschaft für Psychologie (DGPs). Berufsethische Richtlinien des Berufsverbandes Deutscher Psychologinnen und Psychologen e.V. und der Deutschen Gesellschaft für Psychologie e.V. (zugleich Berufsordnung des Berufsverbandes Deutscher Psychologinnen und Psychologen e.V.) in der von der Delegiertenkonferenz des BDP am 4.6.2016 und von der Mitgliederversammlung der DGPs am 21.9.2016 beschlossenen Fassung.
https://www.dgps.de/fileadmin/user_upload/PDF/berufsethik-foederation-2016.pdf

7.3 Grundsätze für Forschung und Publikation

	beginnen erst mit dem Forschungsprojekt, nachdem sie eine Bewilligung erhalten haben. Sie führen ihr Forschungsprojekt in Übereinstimmung mit dem bewilligten Vorgehen durch.
(3)	<p><i>Auf Aufklärung basierende Einwilligung in die Forschung</i></p> <p>(a) Voraussetzung dafür, dass Psychologinnen und Psychologen persönlich, auf elektronischem Weg oder mit Hilfe anderer Kommunikationsformen Forschung durchführen, ist die persönliche Einwilligung der an der Forschung teilnehmenden Personen. Solche Einwilligungserklärungen basieren stets auf einer Aufklärung über das Forschungsvorhaben, die in verständlicher Form dargeboten wird. Hiervon ausgenommen sind solche Forschungsarbeiten, deren Durchführung durch andere Regelungen in diesen Richtlinien gedeckt ist.</p> <p>(b) Psychologinnen und Psychologen müssen Personen, die von Rechts wegen nicht in der Lage sind, eine auf Aufklärung basierende Einwilligung abzugeben, dennoch (1) ihre Forschungsarbeiten angemessen erklären, (2) um deren individuelles Einverständnis nachsuchen, (3) die Prioritäten und Interessen solcher Personen berücksichtigen und (4) sich die entsprechende Genehmigung einer bevollmächtigten Person verschaffen, wenn eine solche stellvertretende Einwilligung vom Gesetz her vorgeschrieben ist. Wenn die Einwilligung einer bevollmächtigten Person vom Gesetz her nicht vorgeschrieben ist, unternehmen Psychologen geeignete Schritte, um die Rechte und das Wohlergehen des Individuums zu schützen.</p> <p>(c) Psychologinnen und Psychologen dokumentieren in angemessener Weise die schriftliche oder mündliche Einwilligung, die Genehmigung und das Einverständnis.</p> <p>(d) Beim Einholen der auf Aufklärung basierenden Einwilligung klären Psychologinnen und Psychologen die teilnehmenden Personen über folgende Sachverhalte auf: (1) den Zweck der Forschung; (2) die erwartete Dauer der Untersuchung und das Vorgehen; (3) ihr Recht darauf, die Teilnahme abzulehnen oder sie zu beenden, auch wenn die Untersuchung schon begonnen hat; (4) absehbare Konsequenzen der Nichtteilnahme oder der vorzeitigen Beendigung der Teilnahme; (5) absehbare Faktoren, von denen man vernünftigerweise erwarten kann, dass sie die Teilnahmebereitschaft beeinflussen, wie z. B. potenzielle Risiken, Unbehagen oder mögliche anderweitige negative Auswirkungen, die über alltägliche Befindlichkeitsschwankungen hinausgehen; (6) den voraussichtlichen Erkenntnisgewinn durch die Forschungsarbeit; (7) die Gewährleistung von Vertraulichkeit und Anonymität sowie ggf. deren Grenzen; (8) Bonus für die Teilnahme und (9) an wen sie sich mit Fragen zum Forschungsvorhaben und zu ihren Rechten als Forschungsteilnehmerinnen und Forschungsteilnehmer wenden können. Den potenziellen Teilnehmerinnen und Teilnehmern wird die Gelegenheit gegeben, Antworten auf ihre Fragen zum Forschungsvorhaben zu erhalten.</p> <p>(e) Werden Forschungsarbeiten durchgeführt, die Interventionen mit experimentellem Charakter umfassen, werden die teilnehmenden Personen zu Beginn der Forschungsarbeit über Folgendes aufgeklärt: (1) den experimentellen Charakter der Intervention; (2) falls relevant, welche Angebote oder Dienste der Kontrollgruppe zur Verfügung stehen bzw. nicht zur Verfügung stehen; (3) die Kriterien, nach denen die Teilnehmerinnen und Teilnehmer den Experimental- bzw. den Kontrollgruppen zugeordnet werden; (4) verfügbare alternative Interventionen, falls potenziell Teilnehmende nicht an der Forschungsarbeit mitwirken oder die Teilnahme vorzeitig beenden möchten, und (5), falls relevant, wer die Kosten für die durchgeführten Interventionen trägt und ob ggf. diese Kosten von den teilnehmenden Personen getragen werden oder von dritter Seite zu erstatten sind.</p>
(4)	<p><i>Auf Aufklärung basierende Einwilligung für das Aufnehmen von Stimmen oder Bildern im Rahmen eines Forschungsvorhabens</i></p> <p>Psychologinnen und Psychologen holen von den an einer Untersuchung teilnehmenden Personen eine auf Aufklärung basierende Einwilligung ein, bevor sie deren Stimmen aufnehmen oder Bilder aufzeichnen, außer (1) die Forschung umfasst nur die Beobachtung</p>

7.3 Grundsätze für Forschung und Publikation

	<p>natürlichen Verhaltens im öffentlichen Raum und es ist nicht zu erwarten, dass die Aufnahme so genutzt wird, dass eine Person identifiziert wird oder Schaden nimmt; (2) das Forschungsdesign schließt Täuschung ein und die Einwilligung für die Nutzung der Aufnahmen wird im Rahmen der anschließenden Aufklärung erbeten.</p>
(5)	<p><i>Klientinnen und Klienten/Patientinnen und Patienten, Schülerinnen und Schüler, Studierende sowie Psychologinnen und Psychologen unterstellte Personen als Forschungsteilnehmerinnen und -teilnehmer</i></p> <p>(a) Wenn Forschungsarbeiten mit den oben genannten Personen durchgeführt werden, tragen Psychologinnen und Psychologen dafür Sorge, dass eine Nichtteilnahme oder die vorzeitige Beendigung der Teilnahme für die potenziell Teilnehmenden keine nachteiligen Konsequenzen haben wird.</p> <p>(b) Ist die Teilnahme an Forschungsprojekten und Untersuchungen Teil der Ausbildung oder durch Prüfungsordnungen vorgeschrieben, so müssen die potenziell Teilnehmenden auf gleichwertige Alternativen zur Untersuchungsteilnahme hingewiesen werden.</p>
(6)	<p><i>Verzicht auf eine auf Aufklärung basierende Einwilligung in die Forschung</i></p> <p>Psychologinnen und Psychologen können auf eine auf Aufklärung basierende Einwilligung nur dann verzichten, (1) wenn vernünftigerweise davon ausgegangen werden kann, dass die Teilnahme an der Forschung keinen Schaden oder kein Unbehagen erzeugt, die über alltägliche Erfahrungen hinausgehen, und wenn die Forschung sich (a) auf gängige Erziehungsmethoden, Curricula oder Unterrichtsmethoden im Bildungsbereich bezieht; sich (b) auf anonyme Fragen/Fragebögen, freie Beobachtungen oder Archivmaterial bezieht, dessen Enthüllung die teilnehmenden Personen nicht den Risiken einer straf- oder zivilrechtlichen Haftbarkeit, finanzieller Verluste, beruflicher Nachteile oder Rufschädigungen aussetzt und bei denen die Vertraulichkeit gewährleistet ist; sich (c) auf Faktoren bezieht, welche die Arbeits- und Organisationseffizienz in Organisationen betreffen, deren Untersuchung keine beruflichen Nachteile für die teilnehmenden Personen haben können und bei denen die Vertraulichkeit gewährleistet ist, oder (2) wenn die Forschung anderweitig durch Gesetze und Verordnungen erlaubt ist.</p>
(7)	<p><i>Anreize zur Teilnahme an Forschungsvorhaben</i></p> <p>(a) Psychologinnen und Psychologen vermeiden übertriebene oder unverhältnismäßige finanzielle oder anderweitige Anreize bei der Anwerbung von an der Forschung teilnehmenden Personen, wenn anzunehmen ist, dass solche Anreize zu einer Teilnahme nötigen würden.</p> <p>(b) Wenn berufliche Leistungen oder Dienste (z. B. Therapie, Beratung) als Anreiz zur Teilnahme angeboten werden, erläutern Psychologinnen und Psychologen die Art der Dienstleistung sowie die mit ihnen verbundenen Risiken, Verpflichtungen und Grenzen.</p>
(8)	<p><i>Täuschung in der Forschung</i></p> <p>(a) Psychologinnen und Psychologen führen keine Studie auf der Basis von Täuschung durch, es sei denn, sie sind nach gründlicher Überlegung zu dem Schluss gekommen, dass der Einsatz von Täuschungstechniken durch den voraussichtlichen bedeutsamen wissenschaftlichen, pädagogischen oder praktischen Erkenntnisgewinn gerechtfertigt ist und dass geeignete alternative Vorgehensweisen ohne Täuschung nicht zur Verfügung stehen.</p> <p>(b) Psychologinnen und Psychologen täuschen potenzielle Teilnehmerinnen und Teilnehmer nicht über solche Aspekte einer Forschungsarbeit, von denen vernünftigerweise angenommen werden kann, dass sie ernsthafte physische und/oder psychische Belastungen erzeugen.</p> <p>(c) Psychologinnen und Psychologen klären jede Täuschung innerhalb eines Experiments so früh wie möglich auf, vorzugsweise am Ende der Teilnahme, aber spätestens am Ende der Datenerhebung, und erlauben den teilnehmenden Personen das Zurückziehen ihrer Daten.</p>

7.3 Grundsätze für Forschung und Publikation

(9)	<p><i>Aufklärung der Forschungsteilnehmerinnen und Forschungsteilnehmer</i></p> <p>(a) Psychologinnen und Psychologen informieren die an ihren Untersuchungen Teilnehmenden so bald wie möglich über das Ziel, die Ergebnisse und Schlussfolgerungen aus ihrer Forschungsarbeit, und sie unternehmen geeignete Schritte, um jedes Missverständnis, das teilnehmende Personen haben könnten und das ihnen bewusst ist, zu korrigieren.</p> <p>(b) Wenn wissenschaftliche oder ethische Überlegungen es rechtfertigen, solche Informationen zu verzögern oder zurückzuhalten, ergreifen Psychologinnen und Psychologen geeignete Maßnahmen, um eventuellen Schaden und Risiken abzuwenden bzw. möglichst gering zu halten.</p> <p>(c) Wenn Psychologinnen und Psychologen erfahren, dass Aspekte ihrer Forschung teilnehmenden Personen Schaden zugefügt haben, unternehmen sie geeignete Schritte, um diesen Schaden zu minimieren.</p>
(10)	<p><i>Verantwortungsvoller Umgang mit Tieren in der Forschung</i></p> <p>(a) Der Erwerb und der Umgang mit Tieren zu Forschungszwecken erfolgt von Psychologinnen und Psychologen nur unter Berücksichtigung geltender Gesetze und Verordnungen und in Übereinstimmung mit professionellen Normen und Standards.</p> <p>(b) Alle Forschungsarbeiten mit Tieren werden von Psychologinnen und Psychologen überwacht, die eine Ausbildung in Forschungsmethoden erhalten haben und die Erfahrung im artgerechten Umgang mit Labortieren haben. Sie sind verantwortlich dafür, dass eine angemessene Berücksichtigung des Wohlbefindens, der Gesundheit und der artgerechten Behandlung der Tiere sichergestellt ist.</p> <p>(c) Psychologinnen und Psychologen stellen sicher, dass alle Personen, die unter ihrer Aufsicht mit Tieren arbeiten, Anweisungen bezüglich der Forschungsmethoden und der Pflege, Haltung und Behandlung der Tiere entsprechend ihren Aufgaben erhalten haben.</p> <p>(d) Psychologinnen und Psychologen unternehmen geeignete Anstrengungen, um körperliche Beschwerden, Infektionen, Krankheiten und Schmerzen ihrer Versuchstiere zu minimieren.</p> <p>(e) Psychologinnen und Psychologen wenden Verfahren, die Tiere Schmerzen, Stress oder Entbehrungen aussetzen, nur dann an, wenn alternative Verfahren nicht verfügbar sind und das Forschungsziel durch den zu erwartenden wissenschaftlichen, pädagogischen oder praktischen Erkenntnisgewinn gerechtfertigt ist.</p> <p>(f) Psychologinnen und Psychologen führen Operationen unter angemessener Betäubung durch und wenden Techniken an, um Infektionen zu vermeiden und Schmerzen während und nach der Operation zu minimieren.</p> <p>(g) Sollte es erforderlich sein, das Leben eines Tieres zu beenden, handeln Psychologinnen und Psychologen schnell und mit entsprechend allgemein anerkannten Methoden, um Schmerzen zu minimieren.</p>
(11)	<p><i>Darstellung von Forschungsergebnissen</i></p> <p>(a) Psychologinnen und Psychologen erfinden und fälschen keine Daten.</p> <p>(b) Falls Psychologinnen und Psychologen bedeutsame Fehler in von ihnen veröffentlichten Daten entdecken, unternehmen sie alle Schritte, diese Fehler zu korrigieren, und zwar durch: Berichtigung, Zurückziehen, Erratum oder andere angemessene Publikationsmittel.</p> <p>(c) Daten, die schon früher veröffentlicht wurden, veröffentlichen Psychologinnen und Psychologen nicht als Originaldaten. Dies schließt nicht aus, dass Daten wieder veröffentlicht werden, wenn dies durch einen entsprechenden Hinweis klargestellt wird.</p>
(12)	<p><i>Plagiate</i></p> <p>Psychologinnen und Psychologen präsentieren keine Arbeiten oder Daten anderer als ihre</p>

7.3 Grundsätze für Forschung und Publikation	
	eigenen, auch nicht, wenn diese Quelle zitiert wird.
(13)	<p><i>Kennzeichnung des Leistungsanteils an einer Forschungsarbeit in Publikationen</i></p> <p>(a) Psychologinnen und Psychologen beanspruchen die Verantwortlichkeit für eine Forschungsarbeit, inklusive der Autorenschaft nur dann, wenn sie die Arbeit selbst durchgeführt haben oder maßgeblich daran beteiligt waren.</p> <p>(b) Die Erstautorenschaft oder die Mitautorenschaft spiegeln den Anteil, den eine Autorin bzw. ein Autor an der Forschungsleistung erbracht hat, korrekt wider; sie sind nicht vom beruflichen oder wissenschaftlichen Status der beteiligten Personen beeinflusst. Die bloße berufliche Position, wie z. B. die einer Lehrstuhlinhaberin bzw. eines Lehrstuhlinhabers oder der Leitung einer Forschungseinrichtung, rechtfertigt kein Anrecht auf eine Autorenschaft. Geringe Beiträge zu einer Forschungsarbeit oder zur Erstellung der Publikation werden angemessen gekennzeichnet, z. B. in Fußnoten oder im Vorwort.</p> <p>(c) In der Regel ist eine Doktorandin bzw. ein Doktorand Erstautorin bzw. Erstautor eines von mehreren Autorinnen und/oder Autoren verfassten Artikels, wenn dieser hauptsächlich auf ihrer bzw. seiner Doktorarbeit basiert. Die Betreuerinnen oder Betreuer diskutieren die Autorenschaft für mögliche Publikationen so früh wie möglich und in angemessener Form im Verlauf der Forschung und Publikation mit ihren Doktorandinnen und Doktoranden. Entsprechendes gilt für andere in der Aus-, Fort- und Weiterbildung unter Betreuung entstandene Qualifikationsleistungen.</p>
(14)	<p><i>Weitergabe von Forschungsdaten zum Zweck der Überprüfung</i></p> <p>(a) Nach der Publikation von Daten halten Psychologinnen und Psychologen die Originaldaten nicht zurück, wenn andere Wissenschaftlerinnen und Wissenschaftler ihre Schlussfolgerungen durch Reanalyse überprüfen wollen und die Daten nur zu diesem Zweck nutzen wollen. Dies gilt allerdings nur in solchen Fällen, in denen die Weitergabe der Daten erfolgen kann, ohne dass die Vertraulichkeit personenbezogener Informationen gefährdet ist, und falls keine gesetzlich geregelten Eigentumsrechte der Herausgabe entgegenstehen. Dies impliziert keine kostenlose Weitergabe von Daten; für die Weitergabe der dokumentierten Daten kann der entsprechende Aufwand in Rechnung gestellt werden.</p> <p>(b) Psychologinnen und Psychologen, die von Kolleginnen und Kollegen Daten zur Reanalyse mit dem Ziel der Überprüfung der Schlussfolgerungen anfordern, dürfen diese nur zu dem in der Anfrage angegebenen Zweck verwenden. Für jeden anderweitigen Gebrauch der Daten ist eine vorherige schriftliche Genehmigung einzuholen.</p>
(15)	<p><i>Gutachterinnen und Gutachter</i></p> <p>Psychologinnen und Psychologen, die von anderen eingereichte Unterlagen für Präsentationen, Veröffentlichungen, Stipendien, Anträge auf Drittmittelförderung oder vergleichbare Zwecke begutachten, respektieren die Vertraulichkeit der erhaltenen Informationen und die Eigentumsrechte an diesen Informationen auf Seiten derjenigen, welche diese Unterlagen verfasst haben.</p>
(16)	<p><i>Online-Forschung</i></p> <p>Psychologinnen und Psychologen beachten die Richtlinien 7.3.1 bis 7.3.15 auch bei Online-Forschung. Dies gilt insbesondere für die Regelungen (1) zum Ausschluss von Täuschung bzw. Vorkehrungen, die die Kontaktaufnahme und das nachträgliche Heilen einer Täuschung sicher ermöglichen, (2) zum Ausschluss von Personenidentifikationen (auch bei Aufnahmen im öffentlichen Raum) ohne persönliche Zustimmung der Aufgenommenen sowie (3) zur Sicherheit von Datenerhebungen, -sammlungen und -übermittlungen (insbesondere über Server außerhalb der deutschen Gesetzesregelung).</p>

Forschende können ethischen Bedenken im Zusammenhang mit der Täuschung von Versuchspersonen begegnen, indem sie eine sorgfältige **postexperimentelle Aufklärung** durchführen. Dieser Schritt erfolgt am Ende des Versuchs und umfasst die möglichst vollständige Aufklärung der

Versuchspersonen über Art und Zweck des Experiments sowie über den Grund für die vorgenommene Täuschung. Milgram hat, zu einer Zeit, als die ethischen Richtlinien noch nicht weit entwickelt waren, diesen Schritt nicht besonders sorgfältig beachtet, sondern ließ manche Versuchspersonen noch lange nach dem Versuch im Glauben, dass sie dem Schüler echte Elektroschocks zugefügt hatten (Perry, 2013). Neben der Richtigstellung der täuschenden Aspekte des Experiments sollte der Prozess der postexperimentellen Aufklärung den Versuchspersonen ein Verständnis für den Zweck der Untersuchung vermitteln. Dadurch können sie eine positive Vorstellung von ihrer Rolle im Experiment gewinnen und die Selbstachtung aufrechterhalten, mit der sie das Labor zu Beginn betreten hatten.

Definition Start

Definition

Postexperimentelle Aufklärung (debriefing): Verfahren, bei dem den Versuchspersonen der Zweck des Experiments, an dem sie sich gerade beteiligt haben, erklärt wird; es werden dabei alle Fragen beantwortet, die die Versuchsperson möglicherweise hat.

Definition Stop

2.3.2 Experimentelle Versuchspläne

Questions Start

Wann und warum ist es wichtig, bei einem Experiment eine Kontrollbedingung einzuplanen?

Questions Stop

Wir haben bereits erfahren, wie wichtig es ist, dass die Versuchspersonen den verschiedenen Bedingungen eines Experiments zufällig zugewiesen werden. Ist das nicht der Fall, ist es nur eingeschränkt möglich, die beobachteten Unterschiede, die zwischen den Bedingungen bezüglich der abhängigen Variablen bestehen, auf Veränderungen der unabhängigen Variablen zurückzuführen. Im Folgenden wollen wir uns näher mit der Frage beschäftigen, wie bereits bei der Versuchsplanung (experimental design) Alternativerklärungen so weit wie möglich ausgeschlossen werden können.

Betrachten wir zunächst einen Versuchsplan, der als experimentell *erscheinen* mag, aber nicht als echtes Experimentaldesign gelten kann. Es handelt sich um die sogenannte **One-Shot-Fallstudie** (Cook & Campbell, 1979). Um ein konkretes Beispiel zu nehmen: Angenommen, ein Forschungsteam möchte die Effekte einer neuen Unterrichtsmethode auf das Lernen untersuchen. Sie gehen in eine Schulklasse, führen die neue Methode ein und erfassen, wie gut die Schülerinnen und Schüler das unterrichtete Material verstehen. Welche Schlüsse können sie aus einem solchen Design ziehen? Im strengen Sinne gar keine. Denn es fehlt ein Vergleichswert, um zu beurteilen, ob sich das Verständnis der Schülerinnen und Schüler verbessert oder verschlechtert hat (oder gleich gut geblieben ist).

Definition Start

Definition

One-Shot-Fallstudie (one-shot case study): Forschungsdesign, bei dem Beobachtungen an einer Gruppe gemacht werden, nachdem ein Ereignis vorgefallen ist oder eine Manipulation durchgeführt wurde.

Definition Stop

Wenn wir herausfinden möchten, ob die Intervention irgendeine Wirkung hat, dann besteht eine mögliche Strategie darin, die abhängige Variable (Verständnis) sowohl *vor* als auch *nach* der Intervention zu messen. Es ist dann möglich, zu untersuchen, wie sich die Werte in Bezug auf die gemessene Variable verändern. Ein solches Vorgehen wird als **Messwiederholungsdesign** oder auch

als *Within-Subjects-Design* bezeichnet. Aus der Veränderung der gemessenen Werte alleine können wir jedoch nicht ableiten, dass es sich hierbei um die Wirkung der Intervention handelt. Denn selbst wenn die Verständniswerte der Versuchspersonen nach der Intervention tatsächlich höher sein sollten als zuvor, können wir nicht wissen, ob diese Veränderung nicht auch ohne unser Eingreifen aufgetreten wäre. Mögliche Ursachen hierfür könnten beispielsweise andere Faktoren sein, die zeitgleich zur Intervention gewirkt haben, oder einfach der Umstand, dass die Schülerinnen und Schüler im Laufe der Zeit ohnehin mehr über die Lehrinhalte erfahren haben. Einfach ausgedrückt, können wir somit nicht sicher sein, ob die Intervention die beobachteten Veränderungen *verursacht* hat oder nicht.

Definition Start

Definition

Messwiederholungsdesign: Ein Versuchsplan, in dem die abhängigen Variablen bei denselben Versuchspersonen mehr als einmal gemessen werden.

Definition Stop

Um direkt zu prüfen, ob die Intervention einen Effekt hat, ist es daher unerlässlich, die Verständniswerte der Personen, die die Intervention erhalten haben, mit den Werten derer, die sie *nicht* erhalten haben, vergleichen zu können. Wir müssen dem Design also eine Kontrollbedingung hinzufügen. Die Studie hätte dann zwei Bedingungen, die sich *zwischen den Versuchspersonen* (between-subjects design) unterscheiden, d. h., jede Person in der Studie nimmt nur an einer der beiden Bedingungen teil. In der Experimentalbedingung werden die Versuchspersonen der Manipulation ausgesetzt (die Versuchspersonen in dieser Bedingung stellen die **Experimentalgruppe** dar) und die möglichen Effekte der Manipulation gemessen. In der Kontrollbedingung (die Versuchspersonen in dieser Bedingung stellen die **Kontrollgruppe** dar) wird dieselbe abhängige Variable zum selben Zeitpunkt wie bei der Experimentalgruppe erhoben, ohne dass jedoch zuvor dieselbe Manipulation wie in der Experimentalgruppe durchgeführt worden wäre. Jetzt kann die in der Experimentalbedingung gemachte Beobachtung mit der in der Kontrollbedingung verglichen werden. Also könnten die Forschenden eine Gruppe von Schülerinnen und Schülern, bei der die neue didaktische Methode verwendet wurde, in Bezug auf das Verständnis des gelehrtens Stoffs mit einer anderen Gruppe vergleichen, bei der weiterhin die herkömmliche Methode verwendet wurde. Ein wichtiger Punkt besteht darin, dass die Versuchspersonen den beiden Bedingungen *zufällig* zugewiesen werden. Dadurch wird die Möglichkeit ausgeschlossen, dass die Unterschiede zwischen den Bedingungen auf Unterschieden zwischen den beiden Gruppen von Versuchspersonen beruhen, die bereits vor der Einführung der neuen Unterrichtsmethode vorhanden waren. Wenn sich bei den Schülern und Schülerinnen, die per Zufall der Experimentalbedingung zugewiesen wurden, nach der Intervention ein besseres Verständnis des Lehrstoffs beobachten lässt als bei den Schülern und Schülerinnen der Kontrollgruppe, wurde dieser Unterschied offenbar durch die neue Unterrichtsmethode verursacht.

2.3.2.1 Vor- und Nachteile von Messwiederholungsdesigns

Wie wir bereits besprochen haben, ist es naheliegend, eine Variable vor und nach der Manipulation zu messen, um zu prüfen, ob sich ihre Ausprägung durch die Manipulation verändert hat. Solche Messwiederholungen sind jedoch in der experimentellen sozialpsychologischen Forschung eher selten. Stattdessen werden die Ergebnisvariablen typischerweise erst nach der Manipulation und nicht auch vorher erfasst. Die Werte der Ergebnisvariablen nach der Manipulation werden dann mit den Ergebnissen einer Kontrollbedingung verglichen. Dies wird als **Nur-Nachtest-Kontrollgruppendesign** (post-test only control group design) bezeichnet.

Für den Verzicht auf die Messung der Ergebnisvariablen vor der experimentellen Manipulation gibt es gute methodische Gründe. Diese beziehen sich vor allem auf die sogenannte **Reaktivität**, d. h. auf das Ausmaß, in dem die Messung einer Variablen (oder jeder beliebige Aspekt des experimentellen Vorgehens) die erhobene Variable und ihre nachfolgenden Messungen beeinflussen kann. So könnte

sich beispielsweise die Messung des Verständnisses vor der Manipulation auf verschiedene Weisen auf das Verständnismaß nach der Manipulation auswirken. Die Versuchspersonen könnten etwa die wahre Absicht einer Studie erahnen, obwohl dies nicht erwünscht ist (Sharpe & Whelton, 2016; s. auch *Hinweise aus der experimentellen Situation* in ► Abschn. 2.3.3). Es könnte auch zu Übungseffekten kommen, bei denen die Versuchspersonen nur dann tatsächlich von der Intervention profitieren, wenn sie mit den Verständnisaufgaben vertraut sind. Da solche Effekte durchaus denkbar sind, wäre damit ungewiss, ob die Intervention die gleiche Wirkung haben würde, wenn die Versuchspersonen nicht vorher getestet worden wären. Die Messung im Vorfeld der Intervention könnte auch einen „Anker“ darstellen, auf den sich die Versuchspersonen später beziehen: Sie könnten z. B. unabhängig von der Intervention motiviert sein, sich über die beiden Tests hinweg konsistent zu verhalten (was mögliche Veränderungen verdecken würde) oder sich zwischen den Testungen zu verbessern (was die Veränderung überbetonen würde).

Dennoch müssen diese Nachteile von Designs mit wiederholten Messungen gegen ihre potenziellen Vorteile abgewogen werden. So können Forschende mit solchen Designs beispielsweise beurteilen, wie eine Manipulation oder Intervention die Ausprägung einer abhängigen Variable *verändert*. Die direkte Einschätzung solch einer Veränderung ist in einem Nur-Nachtest-Kontrollgruppendesign nicht möglich. Es ist zudem auch so, dass Designs mit Messwiederholung empfindlicher dafür sind, den Effekt einer Manipulation zu „finden“. Die Gründe hierfür sind komplex; sie beziehen sich aber auf die statistische Power, die wir später in diesem Kapitel mit Blick auf die Replikation von Forschungsergebnissen näher erläutern werden. Einige Fachleute gehen davon aus, dass Messwiederholungseffekte (z. B. der Unterschied zwischen den Werten vor und nach einer Manipulation) daher robuster sind und in späteren Studien häufiger reproduziert werden können als Effekte, die aus Vergleichen zwischen verschiedenen Versuchspersonen resultieren (z. B. der nach einer Intervention festgestellte Unterschied zwischen der Experimental- und der Kontrollgruppe).

Obwohl es wichtig ist, die Vor- und die Nachteile von Designs mit wiederholten Messungen zu berücksichtigen, haben sich sozialpsychologische Experimente dennoch weitgehend darauf konzentriert, Ergebnismessungen über verschiedene Bedingungen hinweg erst *nach* einer Manipulation miteinander zu vergleichen.

Definition Start

Definition

Nur-Nachtest-Kontrollgruppendesign (post-test only control group design): Die Versuchspersonen werden zufällig einer von zwei Gruppen zugeteilt. Eine Gruppe (d. h. die Experimentalgruppe) wird einer Manipulation bzw. Intervention ausgesetzt, die zweite Gruppe (d. h. die Kontrollgruppe) nicht.

Definition Stop

Definition Start

Definition

Experimentalgruppe (experimental group): Gruppe von Versuchspersonen, die der „Experimentalbedingung“ eines Experiments zugewiesen werden.

Definition Stop

Definition Start

Definition

Kontrollgruppe (control group): Gruppe von Versuchspersonen, die nicht der Manipulation bzw. Intervention aus der Experimentalbedingung ausgesetzt sind.

Definition Stop

Definition Start

Definition

Reaktivität: Ein Messverfahren ist reaktiv, wenn durch die Messung verändert wird, was gemessen werden soll.

Definition Stop

2.3.2.2 Komplexere experimentelle Designs

Als Erweiterungen des Nur-Nachtest-Kontrollgruppendesigns gibt es noch anspruchsvollere und komplexere Designs. Jedes davon stellt einen noch umfassenderen Versuch dar, die Möglichkeit auszuschließen, dass beobachtete Unterschiede zwischen den Bedingungen auf etwas anderes als die Manipulation der unabhängigen Variablen zurückgehen (s. Cook & Campbell, 1979). Ein in der Sozialpsychologie sehr häufig verwendetes Design ist das **faktorielle Experiment**; innerhalb derselben Untersuchung werden dabei zwei oder mehrere unabhängige Variablen manipuliert. Ein faktorielles Design enthält alle möglichen Kombinationen von Ausprägungen dieser unabhängigen Variablen. Das einfachste Beispiel eines solchen Designs besteht aus zwei unabhängigen Variablen mit jeweils zwei Ausprägungen. Aus der Kombination der beiden Variablen resultiert der in • Abb. 2.7 beschriebene Versuchsplan. Der wesentliche Vorteil eines faktoriellen Experiments liegt darin, dass wir die separaten und die *kombinierten* Effekte von zwei und mehr unabhängigen Variablen untersuchen können. Die separaten Effekte der unabhängigen Variablen werden als **Haupteffekte** bezeichnet. Wenn der kombinierte Effekt zweier unabhängiger Variablen von der Summe der beiden Haupteffekte abweicht, wird dies als **Interaktionseffekt** (auch: Interaktion, Wechselwirkungseffekt, Wechselwirkung) bezeichnet.

Definition Start

Definition

Faktorielles Experiment (factorial experiment): Ein Experiment, in dem zwei oder mehr unabhängige Variablen innerhalb des gleichen Designs manipuliert werden.

Definition Stop

Definition Start

Definition

Haupteffekt (main effect): Begriff für die separaten Effekte jeder unabhängigen Variablen in einem ► faktoriellen Experiment.

Definition Stop

Definition Start

Definition

Interaktionseffekt (interaction effect): Von einem Interaktionseffekt (auch: Wechselwirkungseffekt, Wechselwirkung) ist dann die Rede, wenn der kombinierte Effekt von zwei (oder mehreren) unabhängigen Variablen in einem faktoriellen Experiment ein Ergebnismuster zeigt, das von der Summe der ► Haupteffekte abweicht.

Definition Stop

Um eine solche Interaktion zu veranschaulichen, betrachten wir die Untersuchung von Petty, Cacioppo und Goldman (1981) zu den Effekten überredender Kommunikationen auf die Einstellungsänderung (für eine detailliertere Beschreibung der Studie s. ► Aus der Forschung: Verarbeitungsmotivation und Einstellungsänderung in ► Kap. 7). Um das „Modell der Elaborationswahrscheinlichkeit“ zu überprüfen – eine Überredungstheorie von Petty und Cacioppo (1986a, ► Kap. 7) –, manipulierten die Forschenden zwei Variablen. Die erste war die Qualität der Argumente, d. h., ob die studentischen Versuchspersonen mit starken oder schwachen Argumenten dazu bewogen werden sollten, eine Verschärfung der Prüfungsbedingungen an ihrer Hochschule zu befürworten. Die zweite Variable war die Relevanz der Kommunikation für die Zielpersonen, d. h., ob sie dachten, dass ihre Universität das verschärfte Prüfungssystem bereits im nächsten Jahr einführen würde, sodass sie persönlich die Auswirkungen zu spüren bekommen würden (hohe Relevanz), oder erst in zehn Jahren, wovon sie persönlich nicht mehr betroffen sein würden (geringe Relevanz). Nach dem Modell der Elaborationswahrscheinlichkeit sollte die Qualität der Argumente einen stärkeren Einfluss auf die Einstellung haben, wenn die Versuchspersonen vom Thema der Botschaft betroffen sind, als wenn dies nicht der Fall ist. • Abb. 2.8 zeigt die wichtigsten Befunde aus der Untersuchung von Petty et al. (1981). Wie in • Abb. 2.8 zu sehen ist, war der Effekt der Qualität der Argumente auf die Einstellungen tatsächlich viel stärker, wenn die Relevanz hoch, als wenn sie gering war. Da es sich bei dem vorhergesagten Effekt um eine Interaktion handelt, erfordert die Überprüfung dieser Vorhersage ein faktorielles Experiment.

Platzhalter Abbildung Start

Abb. 2.7 Faktorielles experimentelles Design mit zwei Faktoren mit jeweils zwei Ausprägungen

Datei:

Bildrechte: [Urheberrecht beim Autor]

Abdruckrechte: Nicht notwendig

Hinweise Verlag/Setzerei: alt: 24802_6_De_2_Fig10_Print.eps

Platzhalter Abbildung Stop

Platzhalter Abbildung Start

Abb. 2.8 Interaktion zwischen der Argumentqualität und der persönlichen Relevanz; es zeigt sich, dass die Qualität der Argumente einen größeren Effekt auf die Einstellungen hatte, wenn die Relevanz hoch war (Nach Petty, Cacioppo & Goldman, 1981)

Datei:

Bildrechte: [Urheberrecht beim Autor]

Abdruckrechte: Nicht notwendig

Hinweise Verlag/Setzerei: alt: 24802_6_De_2_Fig11_Print.eps

Platzhalter Abbildung Stop

Das Ergebnismuster einer Interaktion wird häufig als *Moderatoreffekt* bezeichnet. Bei einem Moderatoreffekt variiert die Stärke der Wirkung einer Variablen (d. h. der Moderatorvariablen) auf eine andere je nach Ausprägung einer dritten Variablen. Mit anderen Worten besteht die Rolle der dritten Variablen darin, darüber zu bestimmen, *wann* (unter welchen Bedingungen) eine unabhängige Variable eine abhängige Variable beeinflusst. In • Abb. 2.9 z. B. wird der Effekt von Variable X auf Variable O durch Variable Z moderiert. In diesem Beispiel hat X nur dann einen Effekt auf O, wenn $Z = 2$ (• Abb. 2.9b). Wenn $Z = 1$ ist, hat X keinen Einfluss auf O (• Abb. 2.9a). Es ist wichtig, den Moderatoreffekt von einer anderen Art statistischer Beziehung zwischen drei oder mehr Variablen zu unterscheiden, nämlich von einem *Mediationseffekt*.

Um das Konzept der Mediation zu verstehen, sollten Sie sich klarmachen, dass sozialpsychologische Phänomene häufig aus *Ketten* von Ereignissen bestehen. Dieser Umstand lässt sich im Kern auf die Frage reduzieren, ob die Variable X die Variable O *direkt* beeinflusst oder ob die Beziehung

zwischen X und O über eine weitere Variable Z vermittelt wird. Anders ausgedrückt: X mag einen Effekt auf O haben, jedoch möglicherweise keinen direkten. Stattdessen wirkt X womöglich auf Z – die *vermittelnde* Variable –, die wiederum einen Effekt auf O hat (daher wirkt X über Z auf O). Forschende befassen sich daher nicht nur damit, *ob* oder *wann* X auf O wirkt. Sie beschäftigen sich auch damit, *wie* (über welchen Prozess) die Wirkung von X auf O zustande kommt.

Platzhalter Abbildung Start

Abb. 2.9a,b Diagramm zur Veranschaulichung des Moderatoreffekts von Variable Z auf den Zusammenhang zwischen den Variablen X und O. **b** X hat nur dann einen Effekt auf O, wenn Z den Wert 2 hat. **a** Wenn Z den Wert 1 annimmt, ist kein Effekt von X auf O vorhanden

Datei:

Bildrechte: [Urheberrecht beim Autor]

Abdruckrechte: Nicht notwendig

Hinweise Verlag/Setzerei: alt: 24802_6_De_2_Fig12_Print.eps

Platzhalter Abbildung Stop

Platzhalter Abbildung Start

Abb. 2.10a,b Diagramm zur Veranschaulichung des Effekts von Variable X auf Variable O, der über Variable Z vermittelt wird. **a** Wenn Z nicht berücksichtigt wird, hat X einen signifikanten *direkten* Effekt auf O. **b** Wenn Z berücksichtigt wird, ist der direkte Effekt von X auf O nicht signifikant. Jedoch hat X einen *indirekten* Effekt auf O, und zwar über Z

Datei:

Bildrechte: [Urheberrecht beim Autor]

Abdruckrechte: Nicht notwendig

Hinweise Verlag/Setzerei: alt: 24802_6_De_2_Fig13_Print.eps

Platzhalter Abbildung Stop

In der modernen sozialpsychologischen Forschung versuchen Forschende daher oft, **Mediatorvariablen** zu erfassen und dann eine Mediationsanalyse durchzuführen. Für die Durchführung von Mediationsanalysen gibt es klare Regeln (s. Judd & Kenny, 1981a, b; Kenny, Kashy & Bolger, 1998; Preacher & Hayes, 2004). Beispielsweise mag ein Experiment den Befund erbringen, dass es einen kausalen Effekt von X auf O gibt. Zusätzlich messen wir jedoch auch Z und finden, dass diese Variable sowohl mit X als auch mit O zusammenhängt. Wir können dann überprüfen, ob der ursprüngliche Effekt von X auf O kleiner wird oder verschwindet, wenn wir die Zusammenhänge zwischen X und Z sowie zwischen Z und O statistisch berücksichtigen. Daraus ließe sich schließen, dass der Effekt von X auf O durch Z vermittelt (d. h. *mediert*) wird (Baron & Kenny, 1986). Dies wird in • Abb. 2.10 veranschaulicht. In • Abb. 2.10a hat X einen direkten positiven Effekt auf O. In • Abb. 2.10b, in der die Zusammenhänge zwischen X und Z sowie zwischen Z und O berücksichtigt werden, wird der direkte Effekt von X auf O nicht signifikant. Dafür ist der *indirekte* Effekt von X auf O über Z positiv und signifikant.

Definition Start

Definition

Mediatorvariable (mediating variable): Eine Variable, die den Zusammenhang zwischen zwei anderen Variablen vermittelt (mediert).

Definition Stop

2.3.3 Gefahren für die Validität in der experimentellen Forschung

Questions Start

Was ist der Unterschied zwischen interner und externer Validität?

Questions Stop

Der Begriff *Validität* bezieht sich im Forschungskontext darauf, wie gerechtfertigt die Schlussfolgerungen sind, die aus den Befunden eines Experiments gezogen werden. Gute experimentelle Forschung versucht, drei Arten von Validität zu maximieren: interne, externe und Konstruktvalidität.

Die **interne Validität** bezieht sich auf die Gültigkeit der Schlussfolgerung, dass eine beobachtete Beziehung zwischen unabhängiger und abhängiger Variable eine *kausale* Beziehung darstellt; sie wird durch Verwendung eines einwandfreien experimentellen Designs gefördert. Wir haben bereits gesehen, wie das Nur-Nachtest-Kontrollgruppendesign der One-Shot-Fallstudie überlegen ist, weil die Versuchsperson den Bedingungen zufällig zugewiesen werden. Ohne Zufallszuweisung können zahlreiche Faktoren die interne Validität gefährden (Brewer, 2000; Cook & Campbell, 1979). Diesen Gefährdungen gemeinsam ist die Möglichkeit, dass die Forschenden auch dann einen Effekt beobachten würden, wenn sie gar keine Manipulation bzw. Intervention durchgeführt hätten. Zum Beispiel ist *Reifung* eine Validitätsgefährdung im oben beschriebenen One-Shot-Fallstudiendesign, weil Reifung bedeutet, dass sich die Leistung der Schülerinnen und Schüler verbessern kann, ohne dass eine andere Lehrmethode eingesetzt wird. Wenn Versuchspersonen nicht zufällig den Bedingungen zugewiesen werden, kann außerdem eine *Selbstselektion* stattfinden, sodass sich die Eigenschaften der Versuchspersonen in den verschiedenen Bedingungen bereits zu Beginn des Experiments systematisch unterscheiden. Diese Unterschiede können dann (statt der unabhängigen Variablen) für die Unterschiede in der abhängigen Variablen verantwortlich sein.

Definition Start

Definition

Interne Validität (internal validity): Bezieht sich auf die Gültigkeit der Schlussfolgerung, dass Veränderungen der unabhängigen Variablen zu Veränderungen der abhängigen Variablen führen.

Definition Stop

Auch wenn wir nach der Beurteilung der internen Validität Grund zur Annahme haben mögen, dass die *beobachtete* Beziehung zwischen X und O kausal interpretiert werden kann, müssen wir uns fragen, wie gut die intendierten theoretischen Konstrukte durch die unabhängigen und abhängigen Variablen operationalisiert wurden, bevor wir Schlussfolgerungen über die *theoretische* Beziehung zwischen X und O treffen können. Die **Konstruktvalidität** bezieht sich auf die Gültigkeit der Annahme, dass unabhängige bzw. abhängige Variablen die Konstrukte, die sie repräsentieren sollen, adäquat erfassen. Wenn unsere Manipulation der unabhängigen Variablen X eine geringe Konstruktvalidität hat (weil z. B. nicht nur eine Variable, sondern mehrere gleichzeitig manipuliert werden), sind wir in Bezug auf unsere Schlussfolgerungen über die kausale Wirkung von X auf O eingeschränkt. Nehmen wir beispielsweise an, dass Milgram in seinem Experiment zum Gehorsam für jede der vier oben beschriebenen Bedingungen eine andere Versuchsleitung eingesetzt hätte, sodass Versuchsleitung 1 die Versuchspersonen der ersten Bedingung angeleitet hätte, Versuchsleitung 2 die der zweiten usw. Obwohl es effizient erscheinen mag, die Arbeit der Versuchsleitung zwischen verschiedenen Personen aufzuteilen, würde diese Vorgehensweise eine ernstzunehmende Gefahr für die Konstruktvalidität des Experiments bedeuten. Die vier Bedingungen würden sich nämlich nicht mehr *ausschließlich* hinsichtlich der Nähe zum „Opfer“ unterscheiden, sondern wären auch jeweils von einer anderen Versuchsleitung durchgeführt worden. Die unterschiedlichen Grade an Gehorsam, die unter den vier Bedingungen beobachtet werden, beruhen dann *vielleicht* auf dem kausalen Einfluss der

unabhängigen Variablen „Nähe“, vielleicht aber *eher* auf dem der unterschiedlichen Versuchsleitenden (oder auf einer Kombination dieser beiden Faktoren). Das Problem liegt darin, dass es eine **experimentelle Konfundierung** der Variablen „räumliche Nähe“ mit einer zweiten Variablen gäbe, und zwar der Person der Versuchsleitung.

Definition Start

Definition

Konstruktvalidität (construct validity): Gültigkeit der Annahme, dass unabhängige und abhängige Variablen die theoretischen Konstrukte, die sie repräsentieren sollen, adäquat operationalisieren.

Definition Stop

Definition Start

Definition

Experimentelle Konfundierung (experimental confounding): Besteht eine unabhängige Variable aus zwei oder mehr potenziell trennbaren Komponenten, ist sie konfundiert. Ist die unabhängige Variable konfundiert, können Forschende keine eindeutigen kausalen Schlussfolgerungen ziehen.

Definition Stop

Selbst wenn die Konstruktvalidität der unabhängigen Variablen zufriedenstellend zu sein scheint, bleibt die Frage, ob die abhängige Variable erfasst, was sie erfassen soll. Neben der Konfundierung gibt es hauptsächlich drei Gefährdungen der Konstruktvalidität in sozialpsychologischen Experimenten: soziale Erwünschtheit, Hinweise aus der experimentellen Situation (demand characteristics) und Effekte von Versuchsleitungserwartungen.

Soziale Erwünschtheit bezieht sich auf die Tatsache, dass Versuchspersonen in der Regel gerne in einem günstigen Licht gesehen werden möchten und daher womöglich abgeneigt sind, ehrlich über etwas zu berichten, was ihnen vielleicht negativ ausgelegt werden könnte. Auch „zensieren“ Versuchspersonen einige ihrer Verhaltensweisen möglicherweise, um nicht negativ beurteilt zu werden. In dem Maße, in dem die verwendeten Messinstrumente durch Effekte sozialer Erwünschtheit beeinflusst werden, lässt sich das betreffende theoretische Konstrukt nicht richtig erfassen. Ein naheliegendes Mittel zur Verringerung des Einflusses sozialer Erwünschtheit besteht darin, den Prozess der Messung nichtreaktiv (unobtrusive) zu gestalten: Wenn die Versuchspersonen nicht wissen, was eigentlich gemessen wird, können sie ihr Verhalten nicht darauf einstellen.

Definition Start

Definition

Soziale Erwünschtheit (social desirability): Der Begriff beschreibt die Tatsache, dass Versuchspersonen gewöhnlich darauf aus sind, in einem positiven Licht gesehen zu werden, und deshalb ihre Antworten bzw. ihr Verhalten so gestalten, dass sie möglichst nicht negativ bewertet werden.

Definition Stop

Alternativ ist es möglich, die individuellen Unterschiede in der Tendenz zu sozial erwünschten Antworten zu erfassen und sie dann statistisch zu kontrollieren. Paulhus (1984, 1991) hat ein Messinstrument entwickelt, das als Balanced Inventory of Desirable Responding (BIDR) bekannt ist. Dabei handelt es sich um einen aus 40 Items bestehenden Selbstbeurteilungsfragebogen zur Erfassung der Tendenz, sozial akzeptable oder erwünschte Antworten zu geben. Er besteht aus zwei Subskalen, wobei die eine *Selbsttäuschung (self-deceptive enhancement)* erfasst und die andere *Fremdtäuschung (impression management)*. Beispiele für Items aus jeder der beiden Subskalen

finden sich in ► Individuelle Unterschiede: Soziale Erwünschtheit. Wenn Testwerte in einem Selbstbeurteilungsinstrument (wie etwa Einstellungen gegenüber einer ethnischen Minderheit) mit den BIDR-Werten korreliert sind, so deutet dies darauf hin, dass die Selbstbeurteilungen in Richtung sozialer Erwünschtheit verzerrt sind.

Allerdings ist es besser, wenn es gelingt, die Tendenz zu sozial erwünschten Antworten komplett auszuschalten (anstatt diese Tendenz zu erfassen und sie statistisch zu kontrollieren). Denn es ist nicht abschließend geklärt, was Skalen wie der BIDR messen.

Der englische Begriff „demand characteristics“ steht für **Hinweise aus der experimentellen Situation** (► Kap. 1), also Aspekte des Experiments, die der Versuchsperson unbeabsichtigt einen Eindruck davon vermitteln, welche Hypothesen Gegenstand der Untersuchung sind. Personen, die wissen, dass sie an einer wissenschaftlichen Studie teilnehmen, haben häufig Vermutungen darüber, was die Versuchsleitung herauszufinden erwartet. Sie versuchen dann möglicherweise, die erwarteten Reaktionen zu zeigen. Wenn die Versuchspersonen versuchen, den Hypothesen der Versuchsleitung gerecht zu werden, reagieren sie also auf **Hinweise aus der experimentellen Situation**. Orne (1962, 1969) schlug verschiedene Methoden vor, wie sich genauer bestimmen lässt, welche Rolle Hinweise aus der experimentellen Situation in einer gegebenen Experimentalsituation spielen. Beispielsweise empfahl er die Durchführung **postexperimenteller Befragungen**, um herauszufinden, was nach Auffassung der Versuchspersonen das Ziel des Experiments war. Während Orne (1962, 1969) die Verwendung eines Interviews vorschlug, das vorzugsweise von einer anderen Person als der Versuchsleitung durchgeführt wird, kann die postexperimentelle Befragung auch in Form von Fragebogen nach der Untersuchung erfolgen. Forschende sollten natürlich alles in ihrer Macht Stehende tun, um die Wirkung von Hinweisen aus der experimentellen Situation möglichst gering zu halten. Dies lässt sich beispielsweise durch die Anwendung **nichtreaktiver Methoden** erreichen, d. h. durch Messungen, die so unauffällig sind, dass die Versuchspersonen sie gar nicht bemerken. Außerdem ist es möglich, den Versuchspersonen zu sagen, das Ziel des Experiments könne erst nach Abschluss der Untersuchung mitgeteilt werden und es sei bis dahin wichtig, dass sie *nicht* versuchen, die Hypothese zu erraten. Schließlich können mit einer **Cover Story** die Versuchspersonen glauben gemacht werden, das Ziel des Experiments sei ein ganz anderes, als es in Wirklichkeit der Fall ist. Der Einsatz einer Cover Story ist ein oft eingesetztes Mittel, um die Effekte von Hinweisen aus der experimentellen Situation abzuschwächen.

Definition Start

Definition

Hinweise aus der experimentellen Situation (demand characteristics): Hinweisreize im Experiment, die der Versuchsperson als Anhaltspunkt dienen, welche Verhaltensweisen von ihr erwartet werden, d. h. Hinweisreize, die zu einer bestimmten Art von Reaktion „auffordern“ (demand).

Definition Stop

Definition Start

Definition

Postexperimentelle Befragung (post-experimental enquiry): Technik zur Aufdeckung der Wirkung von Hinweisen aus der experimentellen Situation. Die Versuchsperson wird nach der Teilnahme an einem Experiment sorgfältig befragt, um Aufschluss über ihre Wahrnehmungen im Hinblick auf den Zweck des Experiments zu gewinnen.

Definition Stop

Definition Start

Definition

Nichtreaktive Methoden (unobtrusive measures, nonreactive measures): Messmethoden, deren Anwendung die Versuchspersonen nicht bemerken und die daher ihr Verhalten nicht beeinflussen können (► Reaktivität).

Definition Stop

Definition Start

Definition

Cover Story (cover story): Eine falsche, aber plausible Information über den Zweck eines Experiments. Dahinter steckt die Absicht, den Einfluss von Hinweisen aus der experimentellen Situation einzuschränken.

Definition Stop

Exkurs Start

Individuelle Unterschiede

Soziale Erwünschtheit

Paulhus (1984, 1991) entwickelte ein Maß, das als Balanced Inventory of Desirable Responding (BIDR) bekannt ist. Auf der Grundlage ausführlicher Vortests und der Vorhersage tatsächlicher Verhaltensunterschiede war Paulhus in der Lage, zwei Subskalen zu bilden, von denen eine die mit Selbsttäuschung verbundene Selbstaufwertung (Selbsttäuschung) und die andere die Eindrucksmanipulation (Fremdtäuschung) erfasst. *Selbsttäuschung* bezieht sich auf die Tendenz, sich selbst in einem günstigen Licht zu sehen, während *Fremdtäuschung* den absichtlichen Versuch bezeichnet, die eigenen Antworten zu manipulieren, um bei anderen einen positiven Eindruck hervorzurufen. Unten finden sich Beispiellitems für die beiden Subskalen:

Selbsttäuschung

- Mein erster Eindruck von Menschen stellt sich gewöhnlich als richtig heraus.
- Ich weiß immer, warum ich etwas mag.
- Ich bereue meine Entscheidungen nie.
- Ich bin ein vollständig rationaler Mensch.

Von Befragten, die solchen Items zustimmen, wird angenommen, dass sie ein positiv verzerrtes Bild von der eigenen Person haben (diese Annahme beruht darauf, dass diese Aussagen bei den meisten Menschen nur mit sehr geringer Wahrscheinlichkeit zutreffen).

Fremdtäuschung

- Ich fluche nie.
- Ich verschleierte meine Fehler nie.
- Ich werfe nie Abfall auf die Straße.
- Ich lese nie schlüpfrige Bücher oder Zeitschriften.

Von Befragten, die solchen Items zustimmen, wird angenommen, dass sie motiviert sind, sich selbst in einem positiven Licht darzustellen (auch hier ist es bei den meisten Menschen wieder unwahrscheinlich, dass solche Aussagen auf sie zutreffen).

Die Skalen eignen sich dazu, Befragte zu entdecken, die sich selbst in einem günstigen Licht präsentieren. Wenn ihre Testwerte auf den Skalen zu Selbst- und Fremdtäuschung mit anderen

Maßen korrelieren, ist es wahrscheinlich, dass solche Maße von sozialer Erwünschtheit beeinflusst sind; und diese Effekte lassen sich mithilfe statistischer Verfahren auspartialisieren (d. h. „herausrechnen“).

Zu einer deutschsprachigen Übertragung des BIDR von Paulhus (1994) s. Musch, Brockhaus und Bröder (2002).

Exkurs Stop

Eine weitere mögliche Gefährdung der Konstruktvalidität beruht auf der Hypothese oder den Erwartungen, die die Versuchsleitung bezüglich der Studienergebnisse hat. Solche Erwartungen können das Verhalten der Versuchsleitung gegenüber Versuchspersonen unbeabsichtigt beeinflussen und dazu führen, dass sich die Versuchspersonen mit größerer Wahrscheinlichkeit hypothesenkonform verhalten. Rosenthal (1966) nannte dies den **Effekt von Versuchsleitungserwartungen** (► Kap. 1). Die Prozesse, die zu den Effekten von Versuchsleitungserwartungen führen, sind komplex, aber nonverbale Kommunikation ist ein wesentlicher Bestandteil. Eine naheliegende Methode zur Verringerung dieser Effekte besteht darin, die Versuchsleitung hinsichtlich der überprüften Hypothese im Ungewissen („blind“) zu lassen oder ihr zumindest vorzuenthalten, welcher experimentellen Bedingung die jeweilige Versuchsperson zugewiesen wurde. Andere Möglichkeiten bestehen darin, die Interaktion zwischen Versuchsleitung und Versuchsperson zu reduzieren und das Experiment so weit wie möglich zu automatisieren. Tatsächlich wird in einem Großteil der gegenwärtigen sozialpsychologischen Forschung das gesamte Experiment einschließlich aller Instruktionen den Versuchspersonen mittels Computer dargeboten. Offensichtlich beschränkt dies die Möglichkeiten der Versuchsleitung, ihre Erwartungen (entweder verbal oder nonverbal) an die Versuchspersonen zu übermitteln.

Definition Start

Definition

Effekte von Versuchsleitungserwartungen (experimenter expectancy effects): Effekte, die von der Versuchsleitung im Verlauf der Interaktion mit den Versuchspersonen unbeabsichtigt hervorgerufen werden. Diese Effekte entwickeln sich aus dem Wissen der Versuchsleitung über die Hypothese, die überprüft werden soll, und lassen die Wahrscheinlichkeit dafür ansteigen, dass sich die Versuchspersonen so verhalten, wie es der Hypothese der Versuchsleitung entspricht.

Definition Stop

Selbst wenn es der Versuchsleitung gelingt, alle genannten Gefährdungen der internen Validität und der Konstruktvalidität abzuwenden, bleibt noch eine wichtige Frage zur Validität zu klären: Inwieweit kann die Kausalbeziehung zwischen X und O über die spezifischen Umstände des Experiments hinaus verallgemeinert werden? Die **externe Validität** bezieht sich auf die Generalisierbarkeit eines Befunds über die spezifischen experimentellen Umstände hinaus, unter denen dieser von den Forschenden beobachtet wurde.

Definition Start

Definition

Externe Validität (external validity): Betrifft die Generalisierbarkeit von Befunden auf andere Situationen und Populationen als die der Untersuchung.

Definition Stop

Ein wichtiges Element der experimentellen Situation ist natürlich die Frage, wer überhaupt an Experimenten teilnimmt. Eine zentrale Kritik an sozialpsychologischen Experimenten (und an den Experimenten anderer psychologischer Disziplinen) geht dahin, dass es sich bei den Versuchspersonen oft um Universitätsstudierende handelt. Eine Schätzung aus dem Jahr 1985

kommt auf 74 % der Studien (Sears, 1986). Generell kommen die Versuchspersonen psychologischer Forschung typischerweise aus den sogenannten „WEIRD“-Gesellschaften. WEIRD ist ein Akronym für western (westlich), educated (gebildet), industrialized (industrialisiert), rich (reich, wohlhabend), democratic (demokratisch). Daher unterscheiden sie sich von der Mehrheit der Weltbevölkerung, auf die Forschende ihre Ergebnisse möglicherweise verallgemeinern möchten (Henrich et al., 2010). Dies wirft die Frage auf, ob bekannte Befunde und Theorien über Studierende und WEIRD-Populationen hinaus Gültigkeit haben. Forschende wie Henrich und Kollegen (2010) argumentieren, dass wenig Grund zur Annahme bestehe, dass die auf solchen Populationen basierenden Erkenntnisse universell seien. Sie gehen vielmehr davon aus, dass erhebliche interkulturelle Unterschiede zu erwarten sind. Als Reaktion auf diese Kritik antworten andere Forschende jedoch, dass Stichproben von Studierenden und WEIRD-Populationen an sich *nicht* unbedingt eine Bedrohung für die Validität der Forschung darstellen. Dies begründen sie damit, dass es das Ziel vieler sozialpsychologischer Studien sei, die einem Phänomen zugrunde liegenden Prozesse zu verstehen (z. B. Einstellungsänderung oder Stereotypisierung) – und nicht, die allgemeine Bevölkerung zu beschreiben (ein Ziel, für das bestimmte Arten der Umfrageforschung viel besser geeignet sind). Nach diesem Argument würden sich die Prozesse, die etwa einem Phänomen wie Einstellungsänderung oder Stereotypisierung zugrunde liegen, zwischen Studierenden und Nichtstudierenden, WEIRD- und anderen Gesellschaftsformen nicht grundlegend unterscheiden.

Die Verallgemeinerung eines Befunds auf andere Personengruppen ist nur ein Aspekt der externen Validität. Genauso könnten wir uns fragen, ob ein Befund auch über Veränderungen des experimentellen Settings hinweg Bestand hat. Weiterhin könnte es sein, dass der Befund anders ausfällt, wenn andere Operationalisierungen der unabhängigen oder der abhängigen Variablen verwendet werden. Solche Fragen können meist nur geklärt werden, wenn mehrere Studien mit derselben Fragestellung durchgeführt werden, die sich in diesen Aspekten unterscheiden. Dies führt uns zu einem noch grundlegenderen Thema, und zwar zu den Replikationsstudien.

2.3.4 Die Wichtigkeit der Replikation in der Forschung

Ein wissenschaftlicher Befund sollte nur dann als experimentell fundiert angesehen werden, wenn ein angemessen konzipiertes Experiment *selten daran scheitert*, diesen auf gegebenem Signifikanzniveau zu bestätigen. (Fisher, 1926, S. 504)

Der wissenschaftlich belangvolle physikalische Effekt kann ja geradezu dadurch definiert werden, dass er sich regelmäßig und von jedem reproduzieren lässt, der die Versuchsanordnung nach Vorschrift aufbaut. (Popper, 1935, S. 17)

Diese beiden Zitate stammen von zwei der wichtigsten Persönlichkeiten in der Praxis und der Philosophie der Wissenschaft: Ronald Fisher und Karl Popper. Ihre Botschaft ist klar: Ziel der Forschung ist es nicht nur, in einzelnen Studien Theorien zu prüfen und statistisch signifikante Ergebnisse zu erhalten. Vielmehr ist es wichtig, Hypothesen und Theorien *wiederholt* zu prüfen. Denn wenn wiederholte Studien jedes Mal (oder meistens) zu ähnlichen Ergebnissen führen, können wir uns relativ sicher sein, dass ein Befund „echt“ ist und unsere Hypothese unterstützt. Der Prozess der Bestätigung von Ergebnissen über mehrere Studien wird als **Replikation** bezeichnet.

Definition Start

Definition

Replikation (replication): Die möglichst originalgetreue wiederholte Durchführung einer Studie (direkte Replikation), oder die Überprüfung derselben Hypothese mit unterschiedlichen Methoden (konzeptuelle Replikation)

Definition Stop

Die hohe Relevanz der Replikation erklärt, warum viele sozialpsychologische Forschungsarbeiten nicht nur über eine einzelne Studie berichten, sondern über mehrere Studien zu derselben Hypothese. Es kommt auch immer häufiger vor, dass Forschungsarbeiten in erster Linie darauf abzielen, die in

früheren Arbeiten berichteten Ergebnisse zu replizieren, anstatt zu versuchen, eine neue, ungeprüfte Hypothese zu prüfen.

Dabei ist es wichtig zu beachten, dass es verschiedene Arten der Replikation gibt, d. h. unterschiedliche Ansätze zur Prüfung, ob ein Befund zuverlässig reproduziert werden kann. Ein Ansatz ist die sogenannte *direkte Replikation*. Hierbei wird eine Studie so genau wie möglich wiederholt, indem das gleiche Design, die gleichen Methoden, Analysen und Stichprobenmerkmale angewandt werden. Die direkte Replikation ist sinnvoll, um festzustellen, ob ein Befund reproduziert werden kann (und zwar unabhängig davon, wer die Studie durchführt), wodurch das Vertrauen in die tatsächliche Existenz des Phänomens zunimmt. Ein weiterer Ansatz ist die sogenannte *konzeptuelle Replikation*, bei der Forschende dieselbe Hypothese testen wie in der ursprünglichen Studie, aber eine andere Methode verwenden (s. auch ► Kap. 1 für eine Diskussion über die konzeptuelle Replikation). Eine konzeptuelle Replikation beinhaltet beispielsweise die Prüfung einer Hypothese mit unterschiedlichen Materialien (z. B. eine andere Manipulation der zentralen unabhängigen Variablen und/oder eine andere Operationalisierung der zentralen abhängigen Variablen), in einer anderen Testsituation (z. B. im Feld anstatt im Labor), mit einer anderen Stichprobe (z. B. statt mit Psychologiestudierenden mit Nichtstudierende) und/oder in einem anderen Kontext (z. B. in einem anderen Land). Die konzeptuelle Replikation gibt Aufschluss darüber, wie robust oder empfindlich ein Befund gegenüber Variationen der Methode ist, die für die Überprüfung angewandt wird. Dies ist wichtig, um die Generalisierbarkeit eines Befunds und der dazugehörigen Theorie zu beurteilen. Unter Umständen ist die konzeptuelle Replikation auch die einzige Möglichkeit der Replikation, weil eine direkte Replikation nicht durchführbar ist. Dies ist beispielsweise dann der Fall, wenn die Manipulation einer unabhängigen Variablen vom historischen oder kulturellen Kontext abhängt.

2.3.4.1 Wie gut lassen sich sozialpsychologische Befunde replizieren?

Die Relevanz von Replikationsstudien wird auch daran deutlich, dass sich – wie zunehmend befürchtet wird – eine Reihe bekannter Befunde aus der sozialpsychologischen Forschung nicht so einfach reproduzieren lassen. Die Rede ist dabei von der *Replikationskrise*. Diener und Biswas-Diener (2018) verfassten diesbezüglich eine sehr zugängliche Diskussion, einschließlich der Hintergründe, Ursachen und möglicher Lösungen (s. auch Sedlmeier & Renkewitz, 2018).

Wie wir in ► Kap. 1 besprochen haben, war der Ausgangspunkt der Replikationskrise eine Reihe hochkarätiger Versuche, Befunde aus der psychologischen Forschung unabhängig zu replizieren. Eine der bekanntesten Versuchsreihen wurde dabei von einem großen Konsortium aus Forschenden unter der Leitung von Brian Nosek durchgeführt. Dieses Team wählte 100 Studien aus führenden psychologischen Zeitschriften aus und versuchte, ihre Ergebnisse zu replizieren. Dabei wurden, sofern dies möglich war, die Methoden der Originalpublikation verwendet. Die eindrucksvollen Ergebnisse dieser groß angelegten Untersuchung wurden 2015 in der Zeitschrift *Science* (OSC, 2015) veröffentlicht: Insgesamt zeigten 36 % der Replikationsstudien einen signifikanten ($p < .05$) Effekt in die gleiche Richtung wie in der ursprünglichen Studie. Für die Studien aus dem Bereich der Sozialpsychologie lag dieser Wert mit 25 % sogar noch niedriger. Mit anderen Worten konnte das Replikationsprojekt nur bei einem Viertel der sozialpsychologischen Studien den ursprünglichen Befund reproduzieren.

Über die Bedeutung dieser eindrucksvollen Ergebnisse wird viel diskutiert. Ein wichtiger Diskussionspunkt betrifft die Gründe, weshalb eine Replikationsstudie die Originalbefunde nicht zu reproduzieren vermag. Diese Gründe können sehr vielfältig sein und legen unterschiedliche Schlussfolgerungen über die Ergebnisse des Replikationsprojekts nahe. Einerseits kann eine direkte Replikationsstudie, die einen angenommenen Effekt nicht reproduziert, darauf hinweisen, dass es sich bei den ursprünglichen Befunden tatsächlich um einen Zufall oder eine Anomalie handelte, oder dass der berichtete Effekt zumindest sehr viel kleiner ist als ursprünglich angenommen. Andererseits ist es ebenso gut denkbar, dass die verwendeten Materialien für die neue Stichprobe ungeeignet sind (► Kap. 1), dass es einen anderen methodischen oder verfahrenstechnischen Unterschied zur ursprünglichen Studie gibt oder dass der Befund nur für einen bestimmten Kontext oder eine bestimmte Population relevant ist.

Unabhängig von Diskussionen über einzelne Befunde steht allerdings fest, dass die quantitative

Sozialpsychologie durch die Frage herausgefordert wird, wie gut sich ihre Ergebnisse reproduzieren lassen, und damit, inwieweit ihre Ergebnisse als zuverlässig angesehen werden können. Für die „Replikationskrise“ wurden viele mögliche Gründe genannt, von denen wir hier auf zwei genauer eingehen: die geringe *statistische Power* vieler Studien und der *Publikationsbias*. *Statistische Power* bezieht sich auf die Fähigkeit einer Studie, einen tatsächlich existierenden Effekt (z. B. den Unterschied zwischen zwei Experimentalgruppen) zu entdecken. Sie wird von verschiedenen Faktoren beeinflusst, wobei der Stichprobenumfang zu den wichtigsten gehört. Einfach ausgedrückt gilt, dass die Stärke eines Effekts ungenau geschätzt wird, wenn die Stichprobe zu klein ist. In Studien, die eine kleine Stichprobe (und damit geringe statistische Power) haben, ist es wahrscheinlicher, dass bei einem signifikanten Ergebnis die Effektstärke überschätzt wird, was bedeutet, dass der „echte“ Effekt schwächer ist, als er in der Untersuchung erscheint. *Publikationsbias* bezieht sich wiederum auf die seit Langem bekannte Tendenz, dass Fachzeitschriften bevorzugt statistisch signifikante Ergebnisse veröffentlichen, die als neu und informativ gelten, wohingegen nicht signifikante Ergebnisse oder solche, die eine Nullhypothese bestätigen, in der Regel unveröffentlicht bleiben (Rosenthal, 1979). Das Problem, das aus dieser Verzerrung folgt, ist beträchtlich: Wenn eine veröffentlichte Studie, die einen statistisch signifikanten Effekt berichtet, tatsächlich nur eine von fünf oder sechs ähnlichen Studien ist, und zwar die einzige, die diesen Effekt gefunden hat, dann handelt es sich bei dem veröffentlichten Befund mit größerer Wahrscheinlichkeit um eine Anomalie.

2.3.4.2 Psychologie und Open Science: Wege für die Zukunft

Die Sozialpsychologie stellt sich den Herausforderungen der Replikation und der Reproduzierbarkeit vor allem im Rahmen der *Open Science Initiative*. Open Science bezieht sich auf Prinzipien und Praktiken, die die Transparenz des Forschungsprozesses erhöhen. Forschende verpflichten sich dabei, so viele Details wie möglich über die verwendeten Datenerhebungsmethoden und statistischen Analysen mitzuteilen und die erhobenen Daten anderen Forschenden zur Verfügung zu stellen. Dies ermöglicht es anderen Forschenden, die Ergebnisse durch wiederholte Analysen der Daten zu reproduzieren und eine direkte oder annähernd direkte Replikation der Studie durchzuführen. Eine weitere Praktik innerhalb der Open Science Initiative ist die **Präregistrierung** (© Abb. 2.11).

Platzhalter Abbildung Start

Abb. 2.11 Dieses Symbol kennzeichnet wissenschaftliche Veröffentlichungen mit präregistrierten Studien. (Blohowiak et al. 2020. CC BY 4.0)

Datei:

Bildrechte: [Urheberrecht beim Autor]

Abdruckrechte: Nicht notwendig

Hinweise Verlag/Setzerei: neues Bild

Platzhalter Abbildung Stop

Sie bezieht sich auf die Erhöhung der Transparenz *vor* der Durchführung einer Studie. Hierbei dokumentieren Forschende offen das geplante Vorgehen bei der Datenerhebung, die zu prüfenden Hypothesen und wie sie beabsichtigen, ihre Daten auszuwerten. Die Präregistrierung ermöglicht es somit, abzugleichen, inwiefern der Bericht über die Ausführung der Studie den präregistrierten Plänen entspricht. Die vollständigste Version der Präregistrierung ist der *registered report*, in dem alle Vorüberlegungen, Hypothesen, Methoden und Analysen einer Studie zur Begutachtung durch andere Fachpersonen bei einer Fachzeitschrift eingereicht werden, und zwar noch *bevor die Studie durchgeführt wird*. Das bedeutet, dass die Studie unabhängig davon bewertet wird, wie die Ergebnisse tatsächlich ausfallen. Die Logik dahinter ist, dass die Veröffentlichung einer Studie dann von ihrer Qualität abhängt und nicht davon, ob signifikante Ergebnisse vorliegen. Auf diese Weise würden etwa gut konzipierte Studien, die keine statistisch signifikanten Effekte finden, eher veröffentlicht. Außerdem existieren Zeitschriften, die offener für die Veröffentlichung aussagekräftiger Null-Ergebnisse aus hochwertigen Studien sind, und Open-Access-Repositorien, in denen Forschende Details zu Studien zur Verfügung stellen können, die keine Effekte gefunden haben. Alle diese Wege helfen dabei, dem oben erwähnten *Publikationsbias* entgegenzuwirken. Zusammengefasst zeigt die Open Science Initiative, dass die Reproduzierbarkeit einiger Ergebnisse der Sozialpsychologie

zwar zu Recht infrage gestellt wird, dass sich die Sozialpsychologie aber auch in vorbildlicher Weise darum bemüht, die Transparenz und die Reproduzierbarkeit der Forschung in der Sozialpsychologie (und darüber hinaus) zu verbessern.

Definition Start

Definition

Präregistrierung (pre-registration): Prozess, bei dem Forschende Informationen über eine Studie frei zugänglich machen, bevor diese durchgeführt wird. Dazu gehören etwa die Hypothesen, das Vorgehen bei der Datenerhebung und die Methoden der Datenauswertung.

Definition Stop

2.3.5 Probleme beim Experiment

Questions Start

Welche Hauptkritikpunkte an der Verwendung von Experimenten in der Sozialpsychologie gibt es?

Questions Stop

Die experimentelle Methode ist zentral für den Status der Sozialpsychologie als einer empirischen Wissenschaft. Dennoch gibt es zahlreiche kritische Anmerkungen zur experimentellen Forschung in der Sozialpsychologie. Einige davon beziehen sich auf den Erkenntnisgewinn, andere hinterfragen die Annahmen, die dem Experimentieren zugrunde liegen. Zu diesen Annahmen – die auch von anderen wissenschaftlichen Disziplinen geteilt werden – gehört, dass der untersuchte Gegenstand unabhängig von den Forschenden existiert und dass diese einen „objektiven“ und „neutralen“ Standpunkt einnehmen, wenn es um die Untersuchung des Gegenstands und die Theoriebildung geht. Tatsächlich lehnen einige Forschende – wie etwa jene, die einen prägenden Einfluss auf die oben erörterte Diskursanalyse hatten – diese Annahmen vollständig ab (z. B. Potter & Wetherell, 1987). Andere nehmen einen weniger radikalen Standpunkt ein, hinterfragen aber dennoch die Rolle der Forschenden in Experimentalsituationen und heben hervor, wie der soziale Kontext die Bedeutung und den Wert experimenteller Befunde prägen kann.

Ein Problem betrifft das, was Gergen (1978) als *kulturelle Einbettung* sozialer Ereignisse bezeichnet hat. Damit meint er, dass bei einem Laborexperiment der soziale Kontext mitberücksichtigt werden muss, innerhalb dessen eine unabhängige Variable X manipuliert wurde; denn der Kontext könnte eine Rolle dabei gespielt haben, dass X die beobachteten Effekte auf die abhängige Variable O hervorgerufen hat. Smith, Bond und Kağıtçıbaşı (2006) geben einen Überblick über viele sozialpsychologische Experimente (einschließlich des Milgram-Experiments zum Gehorsam), die in verschiedenen Ländern durchgeführt wurden. Relativ häufig erbrachten diese Experimente in unterschiedlichen kulturellen Kontexten unterschiedliche Befunde (► Kap. 15). Allgemein gesprochen fördert kulturvergleichende Forschung oft zutage, wie vermeintlich universelle psychologische Tendenzen in Wirklichkeit auf kulturellen Normen und Wertvorstellungen beruhen. Dies ist besonders vor dem Hintergrund relevant, dass sich die Sozialpsychologie zumeist auf die oben beschriebenen „WEIRD“-Stichproben stützt. Zum Beispiel untersuchte Miller (1984) die Tendenz, das Verhalten anderer Menschen eher auf dispositionale Faktoren als auf Kontextfaktoren zu attribuieren – den sogenannten fundamentalen Attributionsfehler – sowohl in Amerika als auch in Indien. Sie fand heraus, dass amerikanische Erwachsene in der Tat dazu neigen, das Verhalten anderer durch dispositionale Faktoren zu erklären und weniger durch Faktoren des situationalen Kontexts; im Gegensatz dazu boten indische Erwachsene (Hindus) häufiger Kontexterklärungen an als dispositionale Erklärungen. Laut Miller beruhen diese Unterschiede auf kulturellen Unterschieden der Konstruktion des Selbst im Alltag (für eine Diskussion dieses Themas s. ► Kap. 15).

Ein zweites Problem betrifft die Annahme, dass wir mit Experimenten wissenschaftliche Erkenntnisse über die Prinzipien sozialen Verhaltens gewinnen, die über die Zeit hinweg gültig sind. Es gibt nämlich Gründe, daran zu zweifeln, ob das Experiment (oder irgendeine andere Methode) die empirische

Grundlage für solche Prinzipien liefern kann. Um zu verstehen, warum die Sozialwissenschaften, anders als die Naturwissenschaften, hier ein Problem haben, müssen wir berücksichtigen, dass sich das Verhältnis zwischen den Forschenden und dem Gegenstand der Forschung zwischen den beiden Wissenschaftsbereichen radikal unterscheidet. Im Gegensatz zu den Naturwissenschaften sind in den Sozialwissenschaften die Objekte der Forschung Menschen; Menschen schreiben ihren Handlungen natürlich Bedeutung und Sinnhaftigkeit zu. Die Sozialpsychologie lässt sich daher nicht eindeutig von dem trennen, was sie untersucht; Laien können sich sozialpsychologisches Wissen aneignen und es dazu nutzen, ihre eigenen Handlungen auf eine Weise zu verändern, wie dies Atomen, Elementen und Elementarteilchen nicht möglich ist.

Daraus folgt, dass selbst etablierte sozialpsychologische Theorien nicht als Verkörperung von „Gesetzen“ verstanden werden sollten, die über die Zeit hinweg gültig sind: Wenn Menschen etwas über eine psychologische Theorie erfahren und dadurch veranlasst werden, genau das Verhalten zu ändern, das die Theorie zu erklären versucht, kommt der Theorie offensichtlich nur eine zeitlich begrenzte Gültigkeit zu. An dieser Stelle sei jedoch auch angemerkt, dass manche der Probleme bei der Bewertung von sozialpsychologischen Erkenntnissen durch die Anwendung von **Metaanalysen** gelöst werden können. Bei der Metaanalyse handelt es sich um ein statistisches Verfahren, mit dem die Ergebnisse unabhängiger Studien zum selben Phänomen zusammengeführt werden können, um zu überprüfen, ob ein Befund über diese verschiedenen Studien hinweg zuverlässig festgestellt werden kann (s. Cooper, 1989; Hedges & Olkin, 1985; Johnson & Eagly, 2000). Der immer häufiger zu beobachtende Einsatz der Metaanalyse in der Sozialpsychologie (wenn es relevant erscheint, wird in jedem Kapitel dieses Buchs eine Metaanalyse zitiert) hat zweifellos gezeigt, dass viele sozialpsychologische Behauptungen tatsächlich über mehrere Experimente hinweg bestätigt wurden, die oft aus unterschiedlichen Jahrzehnten stammen. Die Aussagekraft von Metaanalysen kann jedoch durch den *Publikationsbias* eingeschränkt sein, d. h. durch die Tendenz, nur statistisch signifikante Ergebnisse zu veröffentlichen, die wir bereits zuvor in diesem Kapitel besprochen haben (Rothstein, Sutton & Borenstein, 2006). Aus diesem Grund sollte eine gute Metaanalyse versuchen, auch unveröffentlichte und nicht signifikante Ergebnisse zu finden und einzubeziehen. Es liegen zudem auch fortgeschrittene Verfahren vor, um das Vorliegen eines *Publikationsbias* in einer Metaanalyse zu prüfen (z. B. Peters, Sutton, Jones, Abrams & Rushton, 2006).

Definition Start

Definition

Metaanalyse (meta-analysis): Eine Reihe von Techniken, um Ergebnisse unabhängiger Studien zu einem bestimmten Phänomen statistisch zu einem Gesamtergebnis zusammenzuführen. Ziel ist es, herauszufinden, ob sich aus den Befunden über alle Studien hinweg ein zuverlässiges Muster ergibt.

Definition Stop

Bei einer dritten Kontroverse geht es darum, wie sehr die Wertvorstellungen, die Motive und das Verhalten der Versuchsleitung den Forschungsprozess beeinflussen. Dies geht über das hinaus, was wir zuvor über Validität und Hinweise aus der experimentellen Situation gesagt haben. Manche behaupten, dass Forschende nie wirklich „objektiv“ und unparteiisch sein können und dass jeder Aspekt des Forschungsprozesses – einschließlich der Themen, der Theorie und der Hypothesen sowie der Interpretation der Forschungsbefunde – durch das Wertesystem der Forschenden beeinflusst wird. Auch in einem sozialpsychologischen Experiment sind Forschende nur in seltenen Fällen unbeteiligt oder neutral. Vielmehr existiert eine Machtbeziehung zwischen Forschenden und Versuchspersonen, innerhalb derer die Forschenden die Versuchspersonen dazu anleiten, etwas zu tun. Die Versuchspersonen wiederum interpretieren die Handlungen der Forschenden aktiv, ziehen daraus Schlussfolgerungen über die Absichten der Forschenden und reagieren darauf. Dies geschieht oft auf eine komplexe Art und Weise, die die Forschenden möglicherweise nicht erkennen oder bei der sie sich schlicht dafür entscheiden, sie zu ignorieren (Spears & Smith, 2001).

So vehement die experimentelle Methode jedoch kritisiert wurde, so vehement wurde sie auch verteidigt (z. B. Tetlock, 1994). Viele der Kritikerinnen und Kritiker des experimentellen Ansatzes

verlangen zudem nicht, das Experiment völlig aufzugeben. Beispielsweise erkannte Gergen (1973) durchaus an, dass Experimente weiterhin eine wichtige Rolle bei der Aufklärung der Beziehung zwischen biologischen Prozessen (wie etwa physiologischer Erregung) und sozialem Verhalten spielen werden; dass Untersuchungen wie das Milgram-Experiment von Nutzen sind, um die teilweise verhängnisvollen Auswirkungen sozialen Einflusses aufzudecken; dass Experimente den Einfluss von Theorien steigern können, indem sie anschauliche Beispiele für die Bedingungen liefern, unter denen eine Theorie zutreffende Vorhersagen macht; und schließlich, dass Experimente für die Evaluation sozialer Reformen nützlich sein können (beispielsweise für die Einschätzung der Effektivität von Maßnahmen zum Energieeinsparen). Andere (z. B. Spears & Smith, 2001) heben hervor, dass die experimentelle Methode nicht automatisch dadurch infrage gestellt wird, dass wir uns bewusst machen, wie die Versuchsleitung in einem Experiment involviert ist und wie die Versuchspersonen darauf in differenzierter Weise reagieren. Vielmehr erlaubt es uns dieses Vorgehen, die Befunde eines Experiments umfassender und vollständiger zu verstehen.

Results Start

Zusammenfassung

In diesem Abschnitt haben wir verschiedene Aspekte experimenteller Forschung in der Sozialpsychologie betrachtet. Wir haben anfangs die wichtigsten Merkmale des sozialpsychologischen Experiments beschrieben und anschließend einige häufig verwendete experimentelle Designs dargestellt. Danach haben wir die wichtigsten Gefährdungen der Validität experimenteller Forschung betrachtet, z. B. Hinweise aus der experimentellen Situation und den Effekt der Erwartungen der Versuchsleitung. Außerdem haben wir uns mit dem Stellenwert von Replikation und Reproduzierbarkeit beschäftigt. Anschließend wurden diverse mögliche Probleme bei der Verwendung von Experimenten in der sozialpsychologischen Forschung diskutiert.

Results Stop

2.4 Methoden der Datenerhebung

Questions Start

Worin bestehen die Stärken bzw. Schwächen der wichtigsten sozialpsychologischen Datenerhebungsmethoden?

Questions Stop

Unabhängig davon, welcher Forschungsstrategie sich Forschende bedienen, müssen sie entscheiden, mit welcher Methode oder mit welchen Methoden die Daten gesammelt werden sollen. In quantitativen Untersuchungen müssen zur Datenerhebung eine oder mehrere Variablen gemessen werden. So müssen Forschende in korrelativen Studien beispielsweise alle Variablen messen, von denen sie erwarten, dass sie miteinander korrelieren. In experimentellen Designs muss die abhängige Variable gemessen werden. In jedem dieser Fälle muss ein theoretisches Konstrukt (z. B. Aggressivität) in eine messbare Variable überführt werden (z. B. Bereitschaft, jemanden zu verletzen). Jedes Messinstrument in der Psychologie sollte sowohl reliabel als auch valide sein. **Reliabilität** bezieht sich hier auf die Stabilität einer Messung. Angenommen, wir messen die Körpergröße von Erwachsenen, dann wird die Messung von einem auf den nächsten Tag recht stabil bleiben und sie wird unabhängig davon sein, wer die Messung vornimmt. Die Kennzeichen eines reliablen Messinstruments sind: Es ist weder vom Zeitpunkt der Messung noch von der Person abhängig, die die Messung durchführt. Allerdings kann ein Messinstrument hoch reliabel sein und gleichzeitig eine geringe **Validität** aufweisen. Um noch einmal das Beispiel mit der Körpergröße aufzugreifen: Nehmen wir an, wir wollten *eigentlich* das Körpergewicht einer Person messen. Da jedoch keine gut funktionierende Waage, sondern nur ein Metermaß zur Verfügung steht, entscheiden wir uns für die Messung der Körpergröße. Da Körpergröße und Körpergewicht natürlich miteinander korreliert sind, liefert die Messung der Körpergröße sicherlich eine bessere Schätzung des Gewichts, als das Gewicht einfach

zu raten. Es ist jedoch klar, dass die Körpergröße kein besonders valides Messinstrument für das Körpergewicht darstellt. Validität bezieht sich in diesem Zusammenhang darauf, wie gut das Konstrukt, das gemessen werden soll, durch die gemessene Variable erfasst wird.

Definition Start

Definition

Reliabilität (Zuverlässigkeit; reliability): Der Grad, in dem ein Maß frei von Messfehlern ist. Eine Messung ist reliabel, wenn sie bei mehrfacher Durchführung zum gleichen Resultat führt (d. h. zu unterschiedlichen Zeitpunkten oder bei der Erhebung durch unterschiedliche Personen).

Definition Stop

Definition Start

Definition

Validität (Gültigkeit; validity): Eine Messung ist in dem Maße valide, in dem sie genau das misst, was sie zu messen vorgibt.

Definition Stop

In der Sozialpsychologie finden die folgenden Methoden häufige Anwendung: Beobachtungsmaße, Selbstbeurteilungsmaße, implizite und physiologische Maße.

2.4.1 Beobachtungsmaße

Wenn der Gegenstand einer Untersuchung soziales *Verhalten* ist, ist die Beobachtung eine naheliegende Methode. Viele in der Sozialpsychologie interessierende Verhaltensweisen können ohne ausgeklügelte Apparaturen festgestellt werden und manifestieren sich in öffentlichen Situationen; dies macht die betreffenden Verhaltensweisen für Beobachtungen zugänglich. Obwohl Beobachtungsmethoden von informellen und unstrukturierten bis hin zu hoch formalisierten und strukturierten Verfahren variieren können, ist das Ziel in jedem Fall dasselbe: im komplexen Strom sozialen Verhaltens die für die Forschungsfrage möglicherweise relevanten Handlungen zu extrahieren und diese im Verlauf einer gewissen Zeitspanne möglichst vollständig zu registrieren (Weick, 1985).

Gelegentlich erfordert die Forschungssituation oder das Forschungsthema eine informelle und unstrukturierte Beobachtung; dabei sind Forschende unter Umständen gleichzeitig Mitglied der beobachteten Gruppe. Ein klassisches Beispiel für dieses Vorgehen ist die Untersuchung von Festinger, Riecken und Schachter (1956) zu der Frage, wie Menschen damit umgehen, wenn ihre zentralen Überzeugungen klipp und klar widerlegt werden. Die Autoren stießen auf eine religiöse Sekte, die vorhergesagt hatte, dass die nördliche Hemisphäre zu einem bestimmten Zeitpunkt durch eine Flut untergehen werde. Indem Mitglieder des Forschungsteams der Sekte beitraten, konnten sie beobachten, was geschah, als die vorhergesagten Ereignisse nicht eintraten. Unter solchen Umständen muss die Beobachtung natürlich verdeckt und informell erfolgen. Hätten die anderen Sektenmitglieder Verdacht geschöpft, dass die Forschenden keine *echten* Gläubigen waren, hätte die Möglichkeit zur Beobachtung nicht mehr bestanden. Diese Art von Beobachtung wird als **teilnehmende Beobachtung** bezeichnet und liefert im typischen Fall qualitative Daten.

Definition Start

Definition

Teilnehmende Beobachtung (participant observation): Beobachtungsmethode, bei der Forschende die Zielgruppe beobachten, indem sie am Gruppengeschehen teilnehmen, und ihre Beobachtungen sorgfältig aufzeichnen.

Definition Stop

Stärker formalisierte Beobachtungsmethoden können zum Einsatz kommen, wenn das interessierende Verhalten registriert werden kann, ohne dadurch gestört zu werden. So sind beispielsweise im Alltag stattfindende Gespräche oft der Ausgangspunkt von Diskursanalysen und Studien zur Gesprächsanalyse (z. B. die Forschung von Stokoe & Edwards, 2007, über Nachbarschaftsbeschwerden und Polizeiverhöre). Beobachtungen natürlichen Verhaltens können auch quantitativ analysiert werden. Ein Beispiel hierfür ist die Studie von Dixon und Durrheim (2003) über die informelle ethnische Segregation an einem südafrikanischen Strand nach dem Ende der Apartheid. In ihrer Beobachtungsstudie wurden Art und Ausmaß der Segregation grafisch so dargestellt, dass die Verteilung der Mitglieder unterschiedlicher ethnischer Kategorien in unterschiedlichen Bereichen eines Strands bei Durban über die Zeit hinweg erkennbar wurde (• Abb. 2.12). Bei der Analyse der räumlichen Verteilung der Menschen am Strand bemerkten sie, dass Paare und kleine Gruppen von Personen dazu neigten, sich in einer Weise zu gruppieren, die die Forschenden als „Sonnenschirmbereiche“ bezeichneten. Diese Bereiche, in denen sich die typischen Strandaktivitäten entfalteten, wurden oft durch persönliche Besitztümer (z. B. Kleidungsstücke) markiert. Die Forschenden kodierten die ethnische Zusammensetzung von insgesamt 2.654 Sonnenschirmbereichen während einer bestimmten Zeitspanne. Mehr als 99,9 % dieser Bereiche waren ethnisch homogen, entweder ausschließlich „Weiß“ ($N = 1949$; 73,4 %) oder ausschließlich „Schwarz“ ($N = 705$; 26,6 %). Dieses Muster war über die Zeit hinweg stabil; während der gesamten Beobachtungsperiode machten die Forschenden an diesem Strand nur eine einzige ethnisch gemischte Gruppe aus. Dass die Strandbesuchenden im Unwissen darüber gelassen wurden, dass sie an einem Forschungsprojekt teilnehmen, ist durch die zuvor dargestellten ethischen Richtlinien für psychologische Forschung legitimiert (• Tab. 2.1).

Platzhalter Abbildung Start

Abb. 2.12 Verteilung der Besucher und Besucherinnen eines Strandes in der Nähe von Durban (Südafrika) nach ethnischer Zugehörigkeit um 12:30 Uhr am 25. Dezember 1999. Anm.: Jeder rote Punkt entspricht einer Schwarzen Person; ein blauer Punkt entspricht einer Weißen Person, und ein gelber Punkt steht für eine Person mit indischer Abstammung. (Aus: Dixon & Durrheim, 2003. © 2003 The British Psychological Society, mit freundlicher Genehmigung von John Wiley & Sons)

Datei:

Bildrechte: [Urheberrecht beim Autor]

Abdruckrechte: Nicht notwendig

Hinweise Verlag/Setzerei: alt: 24802_6_De_2_Fig14_Print.eps

Platzhalter Abbildung Stop

Eine Gemeinsamkeit dieser beiden Beispiele besteht darin, dass sich die beobachteten Personen der Beobachtung nicht bewusst waren. Eine solchermaßen durchgeführte Beobachtung vermeidet das in der Forschung mit Menschen übliche Problem, dass der Vorgang der Messung als solcher tendenziell einen Einfluss auf das Verhalten der Versuchspersonen hat. Dies ist ein weiteres Beispiel für die zuvor in diesem Kapitel beschriebene Reaktivität. Es ist durch viele Untersuchungen belegt, dass das bloße Wissen darüber, beobachtet zu werden, das gezeigte Verhalten beeinflussen kann (zur Forschung zu den Effekten der sozialen Erleichterung s. ► Kap. 12). Aus diesem Wissen heraus wurden nichtreaktive Methoden zur Beobachtung und Messung von Verhalten entwickelt. Webb, Campbell, Schwartz und Sechrest (2000) haben eine nützliche Quellensammlung zu nichtreaktiven Messmethoden zusammengestellt.

Die am stärksten formalisierte Beobachtungsmethode ist diejenige, bei der die Forschenden ein vorher festgelegtes Kategoriensystem zur Erfassung sozialen Verhaltens verwenden. Auf diese Weise erhobene Daten lassen sich quantitativ auswerten. Ein bekanntes Beispiel für ein solches System ist die Interaktionsprozessanalyse (IPA) von Bales (1950), die zur Untersuchung von Interaktionen in Kleingruppen entwickelt wurde. Hierbei werden die verbalen Interaktionen zwischen den Gruppenmitgliedern im Hinblick auf zwölf vorher festgelegte Kategorien kodiert (z. B. „bittet um Informationen“; ► Kap. 12). Mit den so ermittelten Werten der Gruppenmitglieder kann dann unter

anderem bestimmt werden, wer sich als Führungsperson der Gruppe herauskristallisiert hat (s. Bales & Slater, 1955). Weitere Beispiele für Kodierungsschemata im Zusammenhang mit der Methode der Beobachtung finden sich bei Bakeman (2000).

Beobachtungsmethoden haben im Vergleich zu den im nächsten Abschnitt behandelten Selbstbeurteilungsmaßen zwei grundsätzliche Vorzüge: Erstens können sie häufig nichtreaktiv erhoben werden; zweitens, selbst wenn die Personen wissen, dass sie beobachtet werden, nimmt das von ihnen ausgeführte Verhalten sie zumeist so in Anspruch, dass es weniger leicht ist, ihr Verhalten anzupassen, als die Angaben über ihr Verhalten beim Ausfüllen eines Fragebogens zu beschönigen. Allerdings gibt es Verhaltensweisen, die entweder nur schwer direkt zu beobachten sind (weil sie normalerweise im privaten Bereich stattfinden) oder gar nicht beobachtet werden können (weil sie bereits in der Vergangenheit stattgefunden haben). Darüber hinaus sind wir in der Sozialpsychologie häufig an der Messung von *Wahrnehmungen*, *Kognitionen* oder *Bewertungen* interessiert, die nicht einfach mithilfe von Beobachtungen direkt untersucht werden können. Aus diesen Gründen verwenden Forschende häufig Selbstbeurteilungsmaße.

2.4.2 Selbstbeurteilungsmaße

Ein wesentliches Merkmal der Datenerhebung mithilfe von Selbstbeurteilungsmaßen besteht darin, dass die Versuchsperson selbst über ihre Überzeugungen, Einstellungen, Verhaltensweisen usw. Auskunft gibt. Bei diesen Antworten handelt es sich um Selbstbeurteilungsdaten (self-reports). Selbstbeurteilungsmessungen sind normalerweise schneller, preiswerter und einfacher durchzuführen als Beobachtungsmessungen. Die Forschenden müssen keine Laborsituation realisieren bzw. keine natürliche Situation ausfindig machen, in der sich das Verhalten beobachten lässt. Zudem müssen sie in der Regel keine Beobachtenden schulen oder Aufzeichnungsgeräte verwenden, da Selbstbeurteilungen üblicherweise in schriftlicher Form von den Versuchspersonen selbst vorgenommen werden. Schließlich sind, wie bereits erwähnt, einige der für die Forschenden wichtigsten Variablen nicht direkt beobachtbar. Aus diesen Gründen ist die Erfassung von Selbstbeurteilungen in der sozialpsychologischen Forschung weitverbreitet, und es kommt gar nicht so selten vor, dass Untersuchungen ausschließlich anhand von Selbstbeurteilungen durchgeführt werden.

Es gibt zwei grundlegende Methoden, Selbstbeurteilungsdaten zu sammeln: Fragebogen und Interviews. Bei der Verwendung eines *Fragebogens* erhalten die an einer Studie teilnehmenden Personen eine Reihe von Fragen und zusätzlich Instruktionen dazu, wie sie ihre Antworten schriftlich festhalten sollen (heute geschieht dies oft auch durch internetbasierte Umfragen). Im *Interview* werden den teilnehmenden Personen von einem Interviewer oder einer Interviewerin Fragen gestellt; die Antworten werden dann festgehalten. Interviews sind vor allem von Vorteil, wenn es Grund zur Annahme gibt, dass die Fragen ohne Erklärung möglicherweise schwer verständlich sind. Eine taktvolle und einfühlsame interviewende Person sollte in der Lage sein, eine Beziehung zur interviewten Person aufzubauen und sicherzustellen, dass diese, bevor sie antwortet, die Frage richtig verstanden hat. Ein weiterer Vorteil von Interviews ist die Flexibilität im Hinblick darauf, wie strukturiert sie durchgeführt werden sollen. Strukturierte Interviews folgen einem genau festgelegten Plan in Bezug auf Anzahl, Wortlaut und Reihenfolge der Fragen. Die Befragten können auch dazu aufgefordert werden, die gestellten Fragen mit vorgegebenen Antwortmöglichkeiten wie in einem Fragebogen zu beantworten. Bei halbstrukturierten Interviews werden im Vorfeld festgelegte Themen besprochen, aber die befragende Person kann die Fragen, die sie stellt, an den Gesprächsverlauf anpassen. Interviews können zudem auch relativ unstrukturiert sein, wenn Forschende die Befragten einfach darum bitten, über ein bestimmtes Thema zu sprechen, ohne dass vordefinierte Vorstellungen über die konkret besprochenen Inhalte vorliegen. Diese Methode wird beispielsweise zur Aufzeichnung reichhaltiger biografischer Erzählungen oder „Lebensgeschichten“ eingesetzt (z. B. Hollway & Jefferson, 2005).

Zur Durchführung von Interviews ist jedoch ein erheblicher Aufwand an Zeit und Geld erforderlich; schlecht geschulte Interviewende können die Auskünfte ihres Gegenübers leicht verfälschen, beispielsweise durch das Andeuten einer sozial erwünschten Antwort. Im Gegensatz dazu eignen sich Fragebogen in besonderem Maße dazu, mit minimalen Kosten Daten bei einer großen Zahl von

Personen zu erheben. Die vergleichsweise hohe Anonymität des Vorgangs ist von Vorteil, wenn in den Fragen heikle Themenbereiche angesprochen werden. Allerdings kann es vorkommen, dass viele Personen den Fragebogen nicht oder nur unvollständig ausgefüllt zurücksenden. Die Rücklaufquoten für Fragebogen, die mit der Post an zufällig ausgewählte Namen und Adressen verschickt werden, liegen zwischen 10 und 50 %. Da stets die Gefahr besteht, dass sich die Personen, die nicht antworten, von solchen, die antworten, in bestimmten Aspekten systematisch unterscheiden, sind niedrige Rücklaufquoten unerwünscht.

Die Entwicklung eines guten Fragebogens oder Interviewleitfadens ist schwieriger, als es scheinen mag. Wie bei jedem psychologischen Messinstrument besteht das Ziel darin, Fragen zu formulieren, die reliabel und valide sind. Obwohl es bei der Konstruktion eines Fragebogens viele potenzielle Ursachen mangelnder Reliabilität gibt, besteht die größte Gefahr in der *Mehrdeutigkeit*. Wenn eine Frage mehrdeutig ist, wird sie womöglich von verschiedenen Befragten unterschiedlich interpretiert, und die Antworten beziehen sich dann auf Fragen, die gar nicht in dieser Weise gestellt wurden. Die größte Gefahr für die Validität von Fragen besteht darin, dass sich die Forschenden nicht ausreichend über die spezifischen Ziele im Klaren sind, die sie mit der jeweiligen Frage verfolgen: Je verschwommener die Absicht der Forschenden bei der Formulierung einer bestimmten Frage ist, desto größer ist die Wahrscheinlichkeit, dass diese nicht die Informationen liefert, die für die Untersuchungsziele relevant sind. Es gibt jedoch noch eine Reihe anderer Ursachen für mangelnde Reliabilität und Validität, die schwer zu kontrollieren sind. Eine einfache Faustregel lautet, niemals anzunehmen, dass die Antworten auf eine einzige Frage ein reliables oder valides Messinstrument für ein Konstrukt darstellen. Wenn mehrere Fragen (diese werden auch *Items* genannt) als Messinstrument für ein Konstrukt verwendet werden, sollten sich die verschiedenen Faktoren, die die Reliabilität und die Validität von Antworten auf eine einzelne Frage beeinträchtigen, gegenseitig aufheben. Deshalb stellt im Allgemeinen ein Maß, das auf dem Mittelwert der Antworten auf mehrere Items beruht, ein reliableres Maß für das zugrunde liegende Konstrukt dar.

Es ist schwierig, bei der Konstruktion eines Fragebogens von vornherein alle möglichen Fehlerquellen vorzusehen und auszuschalten. Zur Durchführung von Pilotstudien, bei denen erste Entwürfe des Fragebogens an Gruppen von Befragten getestet werden, gibt es deshalb keine Alternative. Der Fragebogen wird dann aufgrund der Antworten und Kommentare aus den Pilotstudien überarbeitet. Die Konstruktion eines vollständig neuen Fragebogens kann aus diesen Gründen ein zeitaufwendiger und mühsamer Prozess sein. Erfolgreicherweise gibt es Sammlungen bereits entwickelter und vorgetesteter Fragebogen, z. B. die von Robinson, Shaver, Wrightsman und Andrews (1991) oder dem Leibniz-Institut für Sozialwissenschaften (GESIS; <https://www.gesis.org/angebot/>). Im Rahmen der oben besprochenen Open-Science-Praktiken stellen zudem immer mehr Forschende ihre Forschungsmaterialien im Internet zur Verfügung (z. B. auf der Plattform <https://osf.io>). Falls noch kein geeigneter Fragebogen existiert, empfiehlt es sich, vor der Entwicklung eines neuen Fragebogens Texte zur Fragebogenentwicklung (z. B. Moosbrugger & Kelava, 2020; Oppenheim, 1992) zurate zu ziehen.

Selbstbeurteilungsmaße haben verschiedene Vorzüge. Doch worin bestehen ihre Nachteile? Offensichtlich ist es nicht möglich, Selbstbeurteilungen völlig nichtreaktiv zu erheben: Die Befragten sind sich immer der Tatsache bewusst, dass sie untersucht werden, und können ihre Antworten infolgedessen abändern. Insbesondere besteht die Gefahr, dass die Antworten der Befragten durch motivationale Faktoren (z. B. soziale Erwünschtheit) beeinflusst werden. Obwohl es keine einfache Lösung gibt, können Forschende einige Schritte unternehmen, um das Problem einzugrenzen. Erstens ist es wichtig, die Befragten wenn möglich darauf hinzuweisen, dass ihre Antworten anonym sind. Zweitens ist es wichtig, bei Fragen der Bewertung (z. B. der Messung von Einstellungen; ► Kap. 6) zu betonen, dass es keine „richtigen“ oder „falschen“ Antworten gibt. Drittens können Forschende die Motivation der Befragten, unverfälschte Antworten zu geben, dadurch erhöhen, dass sie sie nicht als „Versuchskaninchen“ behandeln, sondern als Kooperationspartnerinnen und -partner. Viertens könnten sie die individuellen Unterschiede in Bezug auf die Tendenz erfassen, sozial erwünschte Antworten zu geben, und sie dann statistisch kontrollieren. Fünftens können Selbstbeurteilungsmaße auch mit der Methode der Beobachtung validiert werden, bei der es ja unwahrscheinlich ist, dass sie denselben Verzerrungen unterliegt, wie sie in Selbstbeurteilungen vorkommen (s. Hewstone, Judd & Sharp, 2011).

2.4.3 Implizite und physiologische Messinstrumente

Eine neuere Entwicklung im Bereich der sozialpsychologischen Forschungsmethoden besteht darin, dass Forschende zunehmend Methoden zur Erfassung von Wahrnehmungen, Kognitionen und Bewertungen einsetzen, die nicht auf Selbstbeurteilungen beruhen und daher nicht die damit verbundenen Nachteile aufweisen. Viele dieser Methoden werden als **implizite Maße** bezeichnet (Greenwald & Banaji, 1995). Obwohl implizite Messinstrumente früher nicht sehr verbreitet waren, haben sie in der Sozialpsychologie durchaus eine lange Tradition: Vor über 70 Jahren veröffentlichte Campbell (1950) einen klassischen Artikel über die indirekte Erfassung von Einstellungen.

Definition Start

Definition

Implizite Maße (implicit measures): Nichtreaktive Messungen von Konstrukten wie Einstellungen (z. B. durch Reaktionszeiten), bei denen die Versuchspersonen nicht wissen, was gemessen wird. Sie gehören zur Gruppe der ► nichtreaktiven Methoden.

Definition Stop

Im Unterschied zu damals werden implizite Maße heute gewöhnlich mithilfe von digitaler Technologie erhoben (Abb. 2.13). Computer, Tablets usw. werden hier nicht nur zur Darbietung des experimentellen Materials eingesetzt, sondern auch – und das ist wichtiger – zur genauen Messung unterschiedlicher Aspekte der Reaktionen von Versuchspersonen auf dieses Material. Ein Beispiel für ein implizites Maß ist die Verwendung von Reaktionslatenzen (d. h., wie lange es dauert, bis eine Versuchsperson auf eine bestimmte Frage antwortet). Solche Maße können neue Einsichten in kognitive Prozesse liefern. Beispielsweise berichtete Mussweiler (2006) über eine Reihe von Studien, bei denen er untersuchte, ob die Ausführung von Bewegungen, die zum Stereotyp einer Gruppe gehören, dazu führt, dass das entsprechende Stereotyp aktiviert wird. Eine frühere Veröffentlichung hatte nahegelegt, dass die Aktivierung des Stereotyps von älteren Menschen Personen dazu veranlasst, langsamer zu gehen (Bargh, Chen & Burrows, 1996 – s. aber auch Doyen, Klein, Pichon & Cleeremans, 2012, für eine andere Sicht auf diesen Effekt). Mussweiler wollte überprüfen, ob auch das Umgekehrte zutrifft. Wenn eine Person langsamer geht, wird dann bei ihr das Stereotyp von älteren Menschen aktiviert? In einer seiner Studien wurden die Versuchspersonen instruiert, mit einer Geschwindigkeit zu gehen, die durch ein Metronom festgelegt war, während sie sich eine über Kopfhörer präsentierte Geschichte anhörten. Die Gehgeschwindigkeit wurde so manipuliert, dass sie entweder 30 Schritte pro Minute betrug (langsam) oder 90 Schritte pro Minute (Kontrollgruppe). Die Versuchspersonen führten dann eine lexikalische Entscheidungsaufgabe durch, bei der sie so schnell und so genau wie möglich beurteilen mussten, ob es sich bei auf dem Computerbildschirm dargebotenen Buchstabenfolgen um Wörter oder um Nichtwörter handelte. Einige der Wörter gehörten zum Stereotyp von älteren Menschen (z. B. *vergesslich*, *weise*) und andere nicht (z. B. *richtig*, *unklar*). Die Versuchspersonen in der Bedingung „langsames Gehen“ erkannten die Wörter, die zum Stereotyp von älteren Menschen gehörten, schneller als die Personen in der Kontrollgruppe; Urteile über die Wörter, die nicht zum Stereotyp gehörten, erfolgten jedoch nicht schneller. Der Vorteil des Einsatzes eines impliziten Messinstruments besteht in diesem Kontext darin, dass sich die Versuchspersonen nicht dessen bewusst sind, was bei der lexikalischen Entscheidungsaufgabe erfasst wird; deswegen können die Befunde nicht im Sinne von Hinweisen aus der experimentellen Situation erklärt werden. Beispiele für den Einsatz impliziter Maße zur Erfassung von Einstellungen finden sich in ► Kap. 6.

Während implizite Maße typischerweise kognitive Prozesse erfassen, werden mit *physiologischen* Methoden biologische Marker oder Korrelate verschiedener psychologischer Zustände gemessen. Ihnen liegt die Logik zugrunde, dass es für Versuchspersonen ähnlich wie bei impliziten Messungen und im Gegensatz zur Methode der Selbstbeurteilung sehr schwierig ist, physiologische Reaktionen vorzutäuschen oder bewusst zu verändern. Tatsächlich hat der Einsatz physiologischer Methoden in der Sozialpsychologie eine lange Tradition, etwa bei der Messung galvanischer Hautreaktionen (ein Maß für physiologische Erregung) als Indikator für rassistische Vorurteile im Kontakt zwischen

unterschiedlichen ethnischen Gruppen (Cooper & Pollock, 1959; Rankin & Campbell, 1955). Dank der erheblichen Fortschritte bei der Entwicklung physiologischer Messtechniken werden sie in sozialpsychologischen Studien immer häufiger eingesetzt. Zum Beispiel wurde in der BBC-Gefängnisstudie von Reicher und Haslam (2006; ► Kap. 8) zur Messung von Stress nicht nur ein Selbstbeurteilungsmaß verwendet, sondern auch Kortisol gemessen – ein Hormon, das unter Stress freigesetzt wird. Ein anderes Beispiel ist die Verwendung kardiovaskulärer Marker als Indikatoren für eine wahrgenommene soziale Bedrohung (Scheepers, De Wit, Ellemers & Sassenberg, 2012). Ein Nachteil physiologischer Methoden ist jedoch, dass sie teuer und technisch anspruchsvoll in ihrer Anwendung sein können.

Platzhalter Abbildung Start

Abb. 2.13 Implizite Messinstrumente stützen sich meist auf Computertechnologie (© Polka Dot RF / Getty Images)

Datei:

Bildrechte: [Urheberrecht beim Autor]

Abdruckrechte: Nicht notwendig

Hinweise Verlag/Setzerei: alt: 24802_6_De_2_Fig15_Print.eps

Platzhalter Abbildung Stop

Ein wichtiger Vorteil impliziter und physiologischer Methoden besteht darin, dass sie nichtreaktiv sind. Das heißt, dass sie keinen Verzerrungen wie etwa sozialer Erwünschtheit oder Hinweisen aus der experimentellen Situation unterliegen. Denn sie erfassen Prozesse, die die Versuchspersonen nicht steuern können (und derer sie sich manchmal gar nicht bewusst sind). Doch daraus folgt nicht automatisch, dass solche Messinstrumente eine hohe Validität besitzen. Spiegelt etwa eine schnelle Reaktionszeit die automatische Wirkung eines Stereotyps wider oder individuelle oder Gruppenunterschiede im lexikalischen Wissen? Um Fragen wie diese zu klären, stehen idealerweise andere Messinstrumente (z. B. Beobachtungsmaße) zur Verfügung, deren Ergebnisse mit der aus den impliziten und/oder physiologischen Maßen gewonnenen Evidenz konvergieren. Prinzipiell trägt solche Evidenz dazu bei, die **konvergente Validität** beider Arten von Maßen sicherzustellen. Konvergente Validität ist gewährleistet, wenn unterschiedliche Operationalisierungen desselben Konstrukts zu den gleichen Ergebnissen führen.

Definition Start

Definition

Konvergente Validität (convergent validity): Ist dann gegeben, wenn zwischen unterschiedlichen Maßen für dasselbe Konstrukt (z. B. Selbstbeurteilung, implizites Maß, Beobachtung) positive Zusammenhänge bestehen.

Definition Stop

Ein weiterer entscheidender Vorteil impliziter und physiologischer Maße besteht darin, dass sie Konstrukte und Prozesse erfassen können, die möglicherweise außerhalb des Bewusstseins der Versuchsperson liegen. Wenn sich Menschen bestimmter Gedanken oder Gefühle nicht bewusst sind, sind sie definitionsgemäß nicht dazu in der Lage, über sie zu berichten, selbst wenn sie hoch motiviert sind, ehrlich zu sein. Die Untersuchung „automatischer“ Prozesse ist zu einem zentralen Thema innerhalb der Forschung zur sozialen Kognition geworden (s. Bargh & Chartrand, 2000; ► Kap. 4). Da für einen automatischen Prozesses unter anderem kennzeichnend ist, dass wir uns dessen nicht bewusst sind, erfordert die Erforschung automatischer Prozesse den Einsatz impliziter Messungen.

2.4.4 Big Data: Soziale Medien und Online-Verhalten als Datenquelle

Die zunehmende Nutzung des Internets in nahezu allen Bereichen des sozialen Lebens und die damit verbundene immense Informationsmenge bietet der Sozialpsychologie eine neue Datenquelle: die sogenannten *Big Data*. Algorithmen können große Mengen an Online-Verhalten registrieren und

zusammenfassen, damit sie anschließend analysiert werden können (z. B. Kosinski, Stillwell & Graepel, 2013). Insbesondere soziale Medien wie Twitter und Facebook sind zu einer sehr nützlichen Datenquelle für Forschende der Sozialpsychologie geworden, die das natürliche Online-Verhalten einer sehr großen Anzahl von Menschen untersuchen möchten. Dabei besteht auch die Möglichkeit, dieses Verhalten im zeitlichen Verlauf abzubilden. Die Themen bisheriger Analysen von Big Data aus sozialen Medien reichen von der Kommunikation zwischen Menschen mit unterschiedlichen politischen Präferenzen (z. B. Barberá et al., 2015) über die Entwicklung sozialer Bewegungen bis hin zur Frage, wie Online-Verhalten, z. B. Twittern, mit Offline-Verhalten, z. B. Protesten, zusammenhängt (Freelon et al., 2018). Obwohl Big Data spannende Forschungsmöglichkeiten bieten, ist dies auch ein Bereich, in dem ethische Forschungsprinzipien sehr sorgfältig geprüft werden müssen (z. B. der Schutz der Privatsphäre der Versuchspersonen), damit sichergestellt werden kann, dass die Daten nicht missbraucht werden.

2.4.5 Wahl der Methode

Alle hier behandelten Datenerhebungsmethoden haben bestimmte Vor- und Nachteile. Obwohl es keine festen Regeln für die Entscheidung gibt, welchem Vorgehen der Vorzug zu geben ist, sind bei der Beurteilung der Eignung eines Verfahrens zwei Dinge zu berücksichtigen. Zum einen lassen sich die unterschiedlichen Arten von Messinstrumenten – Beobachtungen, Selbstbeurteilungen, implizite und physiologische Maße – in vielen Bereichen der Forschung kombiniert einsetzen. Zum anderen unterscheiden sich die Arten von Messinstrumenten hinsichtlich der Informationen, die sie liefern. Wenn Beobachtungs-, Selbstbeurteilungs-, implizite und physiologische Maße desselben Konstrukts auf dieselbe Schlussfolgerung hindeuten, festigt dies das Vertrauen in diese Schlussfolgerung. Des Weiteren lässt sich mithilfe von Selbstbeurteilungsmaßen oft das *Ergebnis* eines Prozesses erfassen; indem Forschende auch Beobachtungsmaße und implizite oder physiologische Maße einsetzen, können sie Einsichten in die *Prozesse* erlangen, durch die das Ergebnis zustande kommt. Eine besondere Eigenschaft impliziter und physiologischer Maße besteht darin, dass Forschende mit ihrer Hilfe Aspekte von Gedanken, Gefühlen und Verhaltensweisen erfassen können, die der betreffenden Person nicht bewusst sind und somit auch nicht verfälscht werden können.

2.4.6 Soziale Neurowissenschaft

Genau genommen ist die **soziale Neurowissenschaft** eher ein interdisziplinäres Forschungsfeld als eine Methode zur Datenerhebung; doch wir reihen sie hier ein, weil Forschende, die den Ansatz der sozialen Neurowissenschaft verfolgen, um Sozialverhalten zu untersuchen, typischerweise neurale Aktivitäten messen und zusätzlich Beobachtungs-, Selbstbeurteilungs- oder implizite Maße verwenden. Obwohl daher zur sozialen Neurowissenschaft viel mehr gehört als nur der Einsatz von Maßen für biologische Prozesse, ist es die Messung derartiger Prozesse, die den Ansatz der sozialen Neurowissenschaft von anderen Ansätzen unterscheidet. Cacioppo et al. (2007, S. 101) formulieren es so:

Eine der sozialen Neurowissenschaft zugrunde liegende Annahme lautet, dass jegliches Sozialverhalten beim Menschen biologisch umgesetzt wird.

Diese Annahme veranlasst die Forschenden, die diesen Ansatz verfolgen, dazu, das „soziale Gehirn“ zu untersuchen. Die beliebteste Methode dafür ist die funktionelle Magnetresonanztomografie (fMRT), eine Technik zur Untersuchung der Hirnregionen, die aktiviert sind, während sich Menschen mit Aufgaben beschäftigen wie der Wahrnehmung des Selbst und anderer Personen, der Regulierung ihrer Emotionen und dem Nachdenken über soziale Gruppen. Mit der fMRT kann der Zustrom sauerstoffreichen Bluts in die aktiven Hirnregionen sichtbar gemacht werden.

Definition Start

Definition

Soziale Neurowissenschaft (social neuroscience): Ein interdisziplinäres Feld mit dem Anspruch, zu verstehen, wie soziale Prozesse und Verhaltensweisen durch biologische Systeme realisiert werden.

Definition Stop

Der Einsatz der fMRT zur Untersuchung des Sozialverhaltens ist eine relativ neue Vorgehensweise und bleibt kostspielig im Hinblick auf die nötige technische Ausrüstung und die für das Training erforderliche Zeit. Dennoch gibt es keinen Zweifel daran, dass ihre Beliebtheit zunehmen wird, wenn Forschende mit den Möglichkeiten, die das Verfahren bietet, vertrauter werden und wenn das Interesse in der Neurowissenschaft wächst, das Sozialverhalten beim Menschen zu untersuchen. Um das Potenzial der fMRT für ein besseres Verständnis sozialpsychologischer Phänomene zu veranschaulichen, verweisen wir auf eine Studie von Harris und Fiske (2006). Sie untersuchten neurale Reaktionen auf „extreme“ Fremdgruppen wie etwa drogenabhängige und obdachlose Menschen, also Gruppen, denen im Sinne des Modells des Stereotypinhalts (► Kap. 14) geringe Werte sowohl im Hinblick auf Wärme als auch im Hinblick auf Kompetenz zugeschrieben werden (im Gegensatz z. B. zu älteren Menschen, die als warm, aber inkompetent eingeschätzt werden, oder reichen Menschen, die als kompetent, aber kalt eingeschätzt werden). Es ist bekannt, dass es ein Areal im Gehirn gibt, den medialen präfrontalen Kortex (mPFC), der bei der Verarbeitung sozialer Informationen aktiviert wird. Harris und Fiske (2006) fanden heraus, dass der mPFC beim Betrachten von Fotos mit Personen aktiviert wird, außer wenn die Person auf dem Foto einer Gruppe angehört, der sowohl in Bezug auf Wärme als auch Kompetenz geringe Werte zugeschrieben werden. In diesem Fall aktiviert das Betrachten der Fotos die Insula, ein Hirnareal, das mit Ekel in Zusammenhang steht. Daher stehen die neuralen Befunde aus dieser Studie im Einklang mit der Auffassung, dass Menschen über extreme Fremdgruppen nicht in derselben Weise denken wie über Mitglieder anderer sozialer Gruppen.

Results Start

Zusammenfassung

Dieser Abschnitt beschäftigte sich mit den wichtigsten Methoden der Datenerhebung, die in der sozialpsychologischen Forschung verwendet werden. Dabei wurden Beobachtung, Selbstbeurteilung, implizite und physiologische Methoden unterschieden. Wir stellten fest, dass jede einzelne Methode ihre Vor- und Nachteile hat und es gute Gründe dafür gibt, mehr als eine Messmethode bei einer Untersuchung einzusetzen. Außerdem behandelten wir die zunehmende Verwendung von Big Data, die aus Online-Verhalten gewonnen werden, sowie die soziale Neurowissenschaft als ein relativ neues interdisziplinäres Forschungsfeld, in dem Forschende Methoden wie bildgebende Verfahren zur Messung der Hirnaktivität einsetzen, um die neuralen Grundlagen des Sozialverhaltens beim Menschen zu untersuchen.

Results Stop

Conclusion Start

Kapitelzusammenfassung

- *Was sind Forschungsmethoden?* Forschungsmethoden sind die Verfahren, mit denen Forschende Informationen sammeln.
- *Wie überprüfen Forschende ihre theoretischen Vorhersagen, und worauf beruhen diese Vorhersagen?* Die Informationen, die mithilfe von Forschungsmethoden gesammelt werden, werden dazu genutzt, die theoretischen Vorhersagen der Forschenden zu überprüfen. Diese Vorhersagen werden aus einer Theorie abgeleitet. Die Theorie entsteht oft durch die Beobachtung von Ereignissen aus dem realen Leben oder durch den Versuch, widersprüchliche Befunde vorangegangener Studien zu klären.
- *Welche Forschungsstrategien gibt es?* Wir haben drei quantitative Forschungsstrategien beschrieben – das Quasiexperiment, das echte Experiment mit Zufallszuweisung und die Umfrageforschung. Sie unterscheiden sich unter anderem darin, (1) wie gut sich die Ergebnisse auf eine Population übertragen lassen, und (2) wie eindeutig die Ergebnisse im Sinne von Kausalzusammenhängen zu interpretieren sind.
- *Was sind qualitative Forschungsansätze?* Hier handelt es sich um Ansätze, die nicht darauf

abzielen, numerische Daten zu liefern. Gewöhnlich konzentrieren sie sich auf die Analyse von Texten und die interpretative Analyse; sie werden oft eingesetzt, wenn Forschende der Auffassung sind, dass quantitative Methoden für die Untersuchung des betreffenden Phänomens nicht angemessen sind. Die Diskursanalyse ist ein beliebter qualitativer Ansatz. Diesem Ansatz liegt die Annahme zugrunde, dass soziale Phänomene im Diskurs konstruiert werden.

- *Was sind die Hauptmerkmale sozialpsychologischer Experimente?* Das Experiment wurde detailliert behandelt, da es sich dabei um die in der Sozialpsychologie vorherrschende Forschungsstrategie handelt. Die Hauptmerkmale experimenteller Forschung sind das experimentelle Szenario, die unabhängige Variable, die abhängige Variable, die Manipulationsüberprüfung und die postexperimentelle Aufklärung.
- *Was ist ein echtes experimentelles Design?* Ein solches Design erlaubt es den Forschenden, den kausalen Schluss zu ziehen, dass Veränderungen der unabhängigen Variablen Veränderungen der abhängigen Variablen bewirken. Ein solches Design muss daher aus mehr als einer Bedingung bestehen, damit die Forschenden Beobachtungen, die unter verschiedenen Bedingungen gemacht werden, miteinander vergleichen können. In einem echten Experiment haben Forschende zudem die Kontrolle darüber, wie die Versuchspersonen auf diese Bedingungen verteilt werden; dies geschieht typischerweise nach dem Zufallsprinzip.
- *Was ist das minimale Design eines echten Experiments?* Dabei handelt es sich um das Nur-Nachtest-Kontrollgruppen-Design, bei dem die Versuchspersonen zufällig einer von zwei Bedingungen zugewiesen werden, von denen nur eine die experimentelle Manipulation beinhaltet. Unter den verschiedenen komplexeren Designs, die zur Verfügung stehen, wird insbesondere das faktorielle Design häufig verwendet, da es die Möglichkeit bietet, Vorhersagen über Interaktionseffekte zu überprüfen.
- *Was bedeutet im Kontext der Forschungsmethoden „Validität“?* Mit Validität ist gemeint, dass die Schlussfolgerungen über eine empirische Studie gerechtfertigt sind. Drei Arten der Validität sind für eindeutige Schlussfolgerungen über sozialpsychologische Forschung nötig: interne, externe und Konstruktvalidität. Wir haben anhand von Beispielen gesehen, in welchen Situationen die jeweiligen Validitäten gefährdet sein können.
- *Warum ist die Replikation in der Forschung so wichtig?* Wir haben hervorgehoben, wie wichtig die Replikation von Ergebnissen über mehrere Studien hinweg ist, bevor sie als „echt“ angesehen werden können. Zudem haben wir zwischen der direkten und der konzeptuellen Replikation unterschieden. Wir haben auch die „Replikationskrise“ in der Psychologie thematisiert, die dadurch entstanden ist, dass einige Versuche gescheitert sind, hochkarätige Ergebnisse unabhängig zu replizieren. Mögliche Wege aus der Krise wurden ebenfalls vorgestellt, insbesondere die Open-Science-Praktiken, die die Transparenz der Forschung erhöhen sollen.
- *Was sind mögliche Nachteile von Experimenten?* Einige Forschende im Bereich der Sozialpsychologie haben die Annahme in Zweifel gezogen, dass sich aus Experimenten so etwas wie empirische „Gesetze“ über menschliches Sozialverhalten ableiten lassen. Von den Argumenten, die gegen diese Annahme sprechen, haben wir die folgenden diskutiert: (1) die kulturelle Einbettung sozialen Verhaltens, (2) die Tatsache, dass soziales Verhalten durch vielerlei Faktoren determiniert wird, und (3) die menschliche Fähigkeit, das eigene Verhalten angesichts sozialpsychologischer Theorien abzuändern.
- *Was sind die grundlegenden Methoden der Datenerhebung in der Sozialpsychologie?* Es wurden einige grundlegende Methoden unterschieden: Beobachtungsmaße, Selbstbeurteilungsmaße, implizite Maße, Big Data und physiologische Maße. Beobachtungsmaße, implizite und physiologische Maße haben den Vorteil, weniger anfällig für Effekte sozialer Erwünschtheit zu sein; sie können darüber hinaus zum Teil völlig nichtreaktiv erhoben werden. Beobachtungsmaße sind jedoch per Definition auf Phänomene beschränkt, die sich beobachten lassen; sie eignen sich nicht zur Erfassung kognitiver Phänomene wie etwa Einstellungen, Kausalattributionen und Stereotype, die direkter Beobachtung nicht zugänglich sind (► Kap. 3, ► Kap. 6, ► Kap. 7 und ► Kap. 14).
- *Wie können solche „verborgenen“ Phänomene untersucht werden, für die sich die Sozialpsychologie interessiert?* Um Phänomene zu untersuchen, die direkter Beobachtung nicht zugänglich sind, haben Forschende traditionell Selbstbeurteilungsmaße eingesetzt. Inzwischen gibt es allerdings eine zunehmende Tendenz, implizite oder sogar physiologische Maße

einzusetzen. Deren Ziel ist es, Phänomene zum Vorschein zu bringen, die entweder außerhalb des individuellen Bewusstseins liegen oder bei herkömmlichen Selbstbeurteilungsmaßen infolge sozialer Erwünschtheit verfälscht wiedergegeben werden.

- Können die unterschiedlichen Methoden der Datensammlung kombiniert eingesetzt werden? Der Vorteil, der damit verbunden ist, dasselbe Konstrukt auf unterschiedliche Arten zu messen, liegt auf der Hand.
- Was haben die neuen bildgebenden Verfahren zur Messung der Hirnaktivität der Sozialpsychologie zu bieten? Forschende der Sozialpsychologie und der Neurowissenschaft werden Techniken wie die bildgebenden Verfahren zur Messung der Hirnaktivität zunehmend nutzen, um die neuralen Grundlagen des menschlichen Sozialverhaltens besser zu verstehen.

Conclusion Stop

DefinitionList Start

Empfohlene Literatur

Aronson, E., Ellsworth, P. C., Carlsmith, J. M., & Gonzales, M. H. (1990). *Methods of research in social psychology* (2nd ed.). New York: McGraw-Hill.

– Eine umfassende Einführung in die Forschungsmethoden der Sozialpsychologie mit einer Betonung der experimentellen Methode.

Cook, T. D. & Campbell, D. T. (1979). *Quasi-Experimentation: Design and analysis issues for field settings*. Chicago: Rand McNally.

– Eine bedeutsame Publikation zu dem Thema, wie die Wahl eines geeigneten Forschungsdesigns dabei hilft, Gefährdungen der Validität auszuschließen.

Fiske, S. T., Gilbert, D. T., & Lindzey, G. (Eds.). (2010). *Handbook of social psychology* (5th ed., Vol. 1). Hoboken, NJ: Wiley.

– Die aktuelle Ausgabe dieses renommierten Handbuchs mit Beiträgen über die experimentelle Methode (Kap. 2), nichtexperimentelle Methoden (Kap. 3) und Auswertung von Daten (Kap. 4).

Greenberg, J., & Folger, R. (1988). *Controversial issues in social research methods*. New York: Springer.

– In diesem Buch werden die Debatten rund um Schlüsselthemen der Forschung sehr gut dargestellt.

Greenwood, J. D. (1989). *Explanation and experiment in social psychological science: Realism and the social constitution of action*. New York: Springer.

– Eine interessante, kritische Erörterung des philosophischen Hintergrunds von Forschungsmethoden.

Haslam, S. A., & McGarty, C. (2003). *Research methods and statistics in psychology*. London: Sage.

– Ein gut lesbares Buch; es ist von Sozialpsychologen verfasst, die der Auffassung sind, dass Psychologie nur verstehen kann, wer den Prozess der Durchführung und der Interpretation psychologischer Forschung versteht.

Reis, H. T., & Judd, C. M. (Eds.). (2014). *Handbook of research methods in social and personality psychology* (2nd ed.). New York: Cambridge University Press.

– Hier werden die wichtigsten methodologischen Themen der Sozial- und der Persönlichkeitspsychologie auf dem neuesten Stand besprochen.

Sedlmeier, P., & Renkewitz, F. (2018). *Forschungsmethoden und Statistik für Psychologen und Sozialwissenschaftler* (3. Aufl.). München: Pearson.

– Dieses Lehrbuch erklärt unter anderem die Logik von Signifikanztests und enthält zwei empfehlenswerte Überblickskapitel zum Thema Replikation und Open Science (Kap. 20 und Kap. 21).

DefinitionList Stop

Abb. 2.1 Anfangsbedingungen, vermittelnder Prozess und Symptome des Gruppendenkens im theoretischen Modell von Janis (1982, © Free Press/Simon & Schuster, © Cengage Learning)

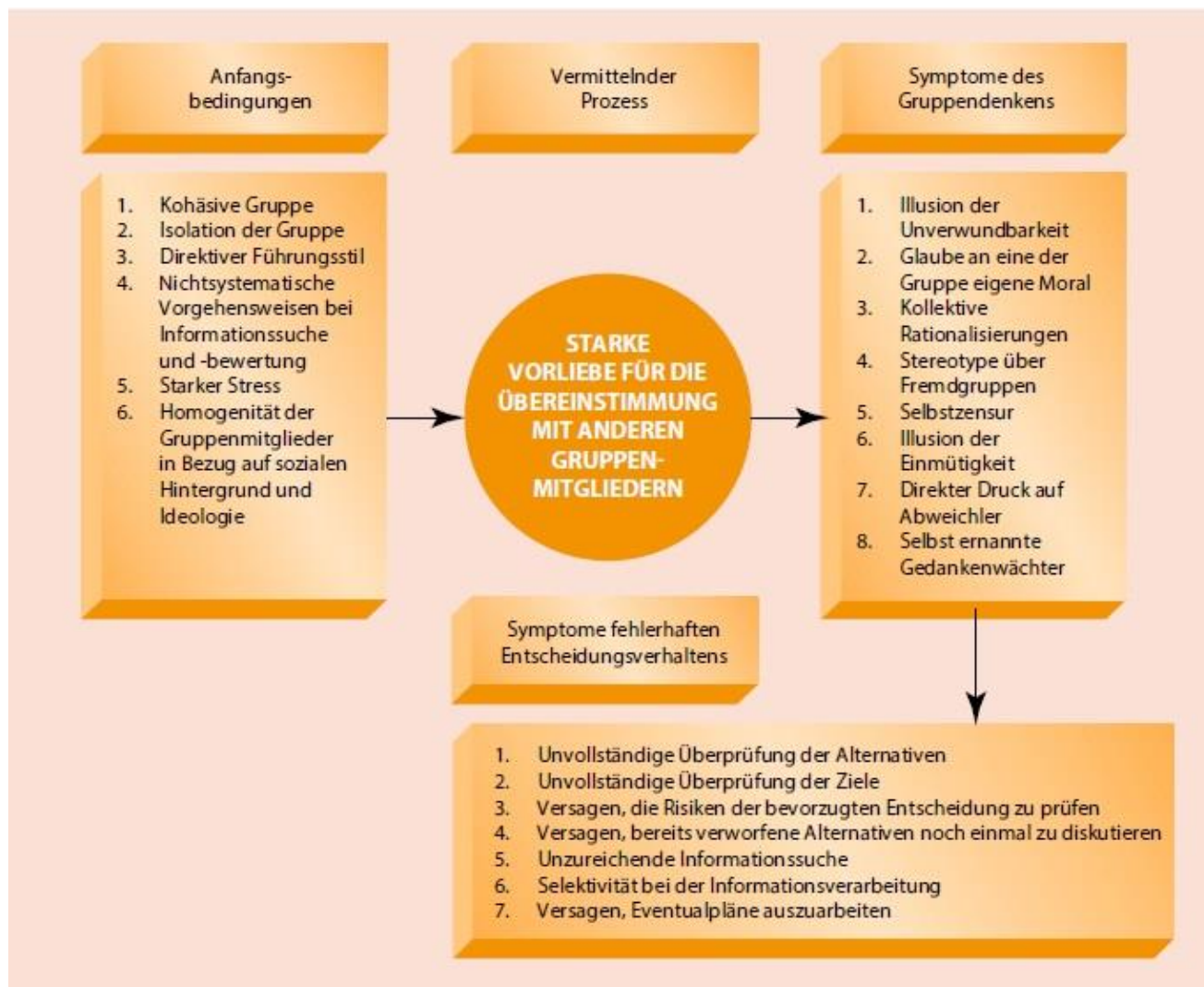


Abb. 2.2 Die selbst beurteilte Neigung, sich aggressiv zu verhalten, in Abhängigkeit von der Art des Films und davon, ob die befragten Personen darauf warteten, einen Film zu sehen, oder ihn gerade gesehen hatten (Daten von Black & Bevan, 1992)

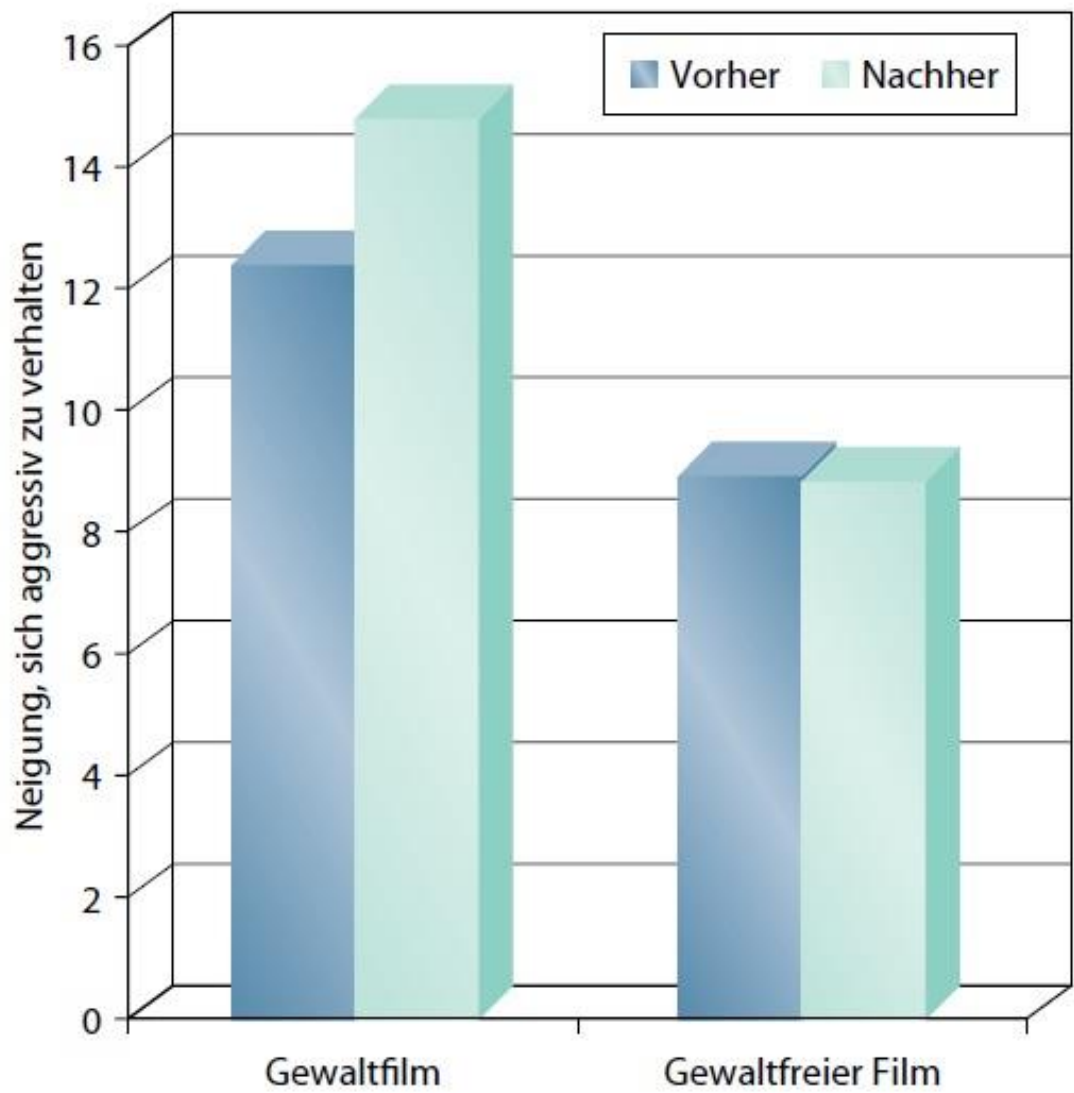


Abb. 2.3 Nimmt bei uns allen die Neigung zu aggressivem Verhalten zu, wenn wir einen gewalthaltigen Film gesehen haben, oder nehmen aggressive Neigungen von vornherein nur bei denjenigen zu, die sich zum Ansehen solcher Filme hingezogen fühlen? (© imago / United Archives)

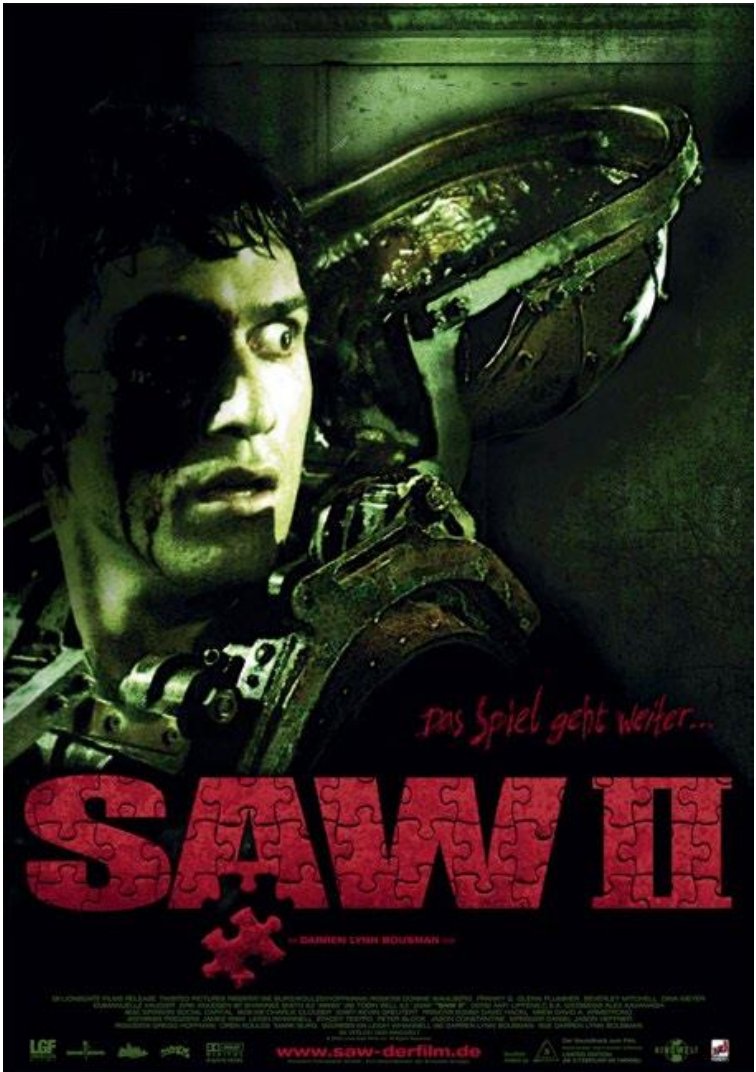


Abb. 2.4 Kann eine Seifenoper dazu beitragen, dass sich die Menschen in Ruanda nach dem Völkermord wieder versöhnen?

noch nicht verfügbar

Abb. 2.5 Eine Strategie zum Sammeln von Forschungsbefunden besteht darin, die öffentliche Meinung mithilfe von Interviews zu erfassen. (© Josef Muellek / Thinkstock)



Abb. 2.6 Prozentsatz der Versuchspersonen, die die maximale Stromstoßstärke anwendeten und die deshalb als völlig gehorsam klassifiziert wurden. (Nach Milgram, 1965)

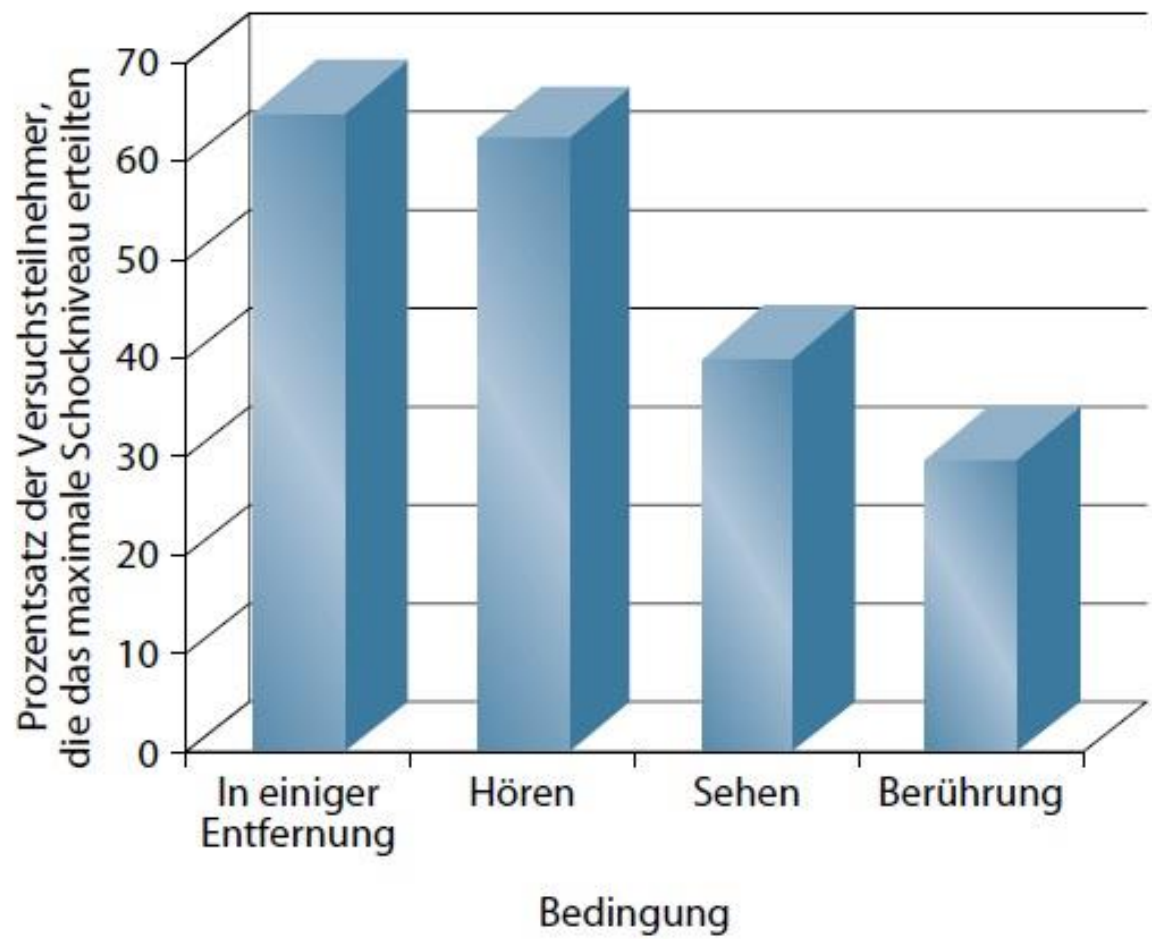


Abb. 2.7 Faktorielles experimentelles Design mit zwei Faktoren mit jeweils zwei Ausprägungen

		Faktor 2	
		Ausprägung 1	Ausprägung 2
Faktor 1	Ausprägung 1		
	Ausprägung 2		

Abb. 2.8 Interaktion zwischen der Argumentqualität und der persönlichen Relevanz; es zeigt sich, dass die Qualität der Argumente einen größeren Effekt auf die Einstellungen hatte, wenn die Relevanz hoch war (Nach Petty, Cacioppo & Goldman, 1981)

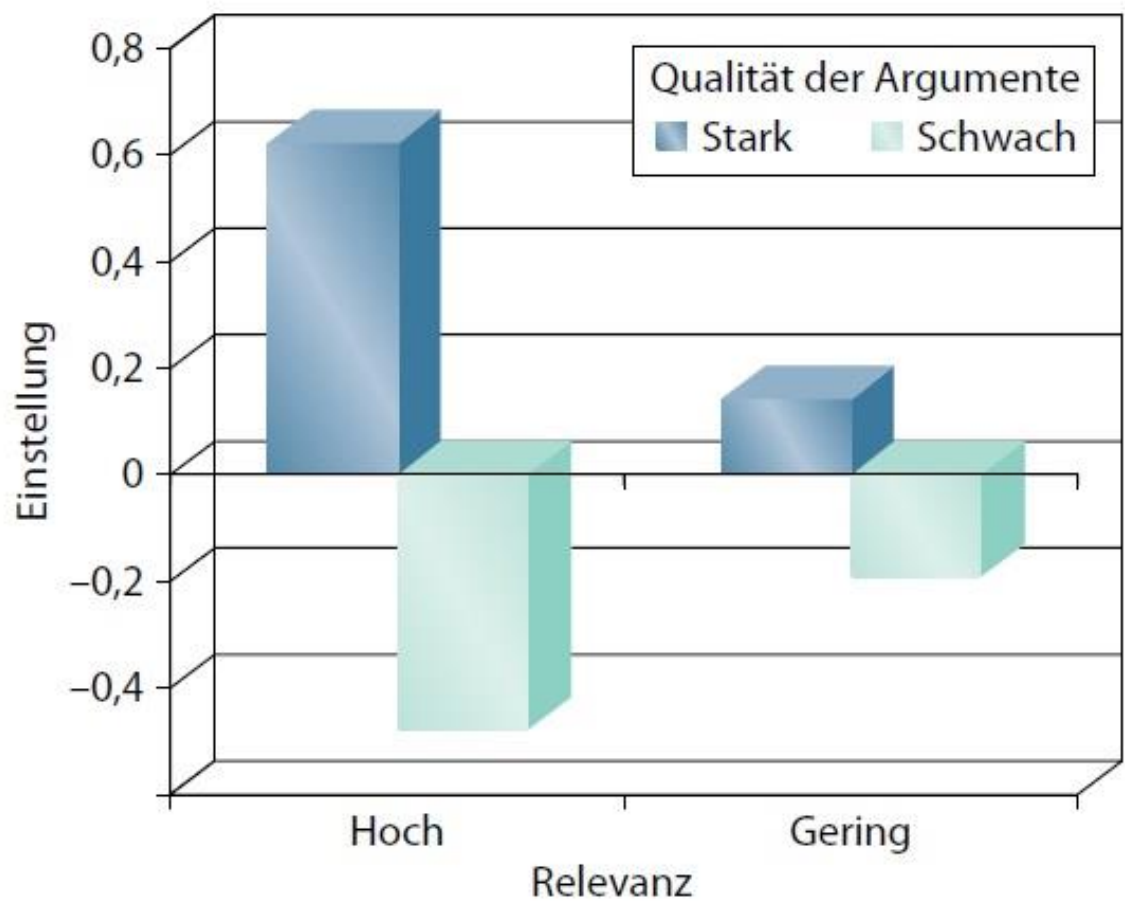


Abb. 2.9a,b Diagramm zur Veranschaulichung des Moderatoreffekts von Variable Z auf den Zusammenhang zwischen den Variablen X und O. b X hat nur dann einen Effekt auf O, wenn Z den Wert 2 hat. a Wenn Z den Wert 1 annimmt, ist kein Effekt von X auf O vorhanden

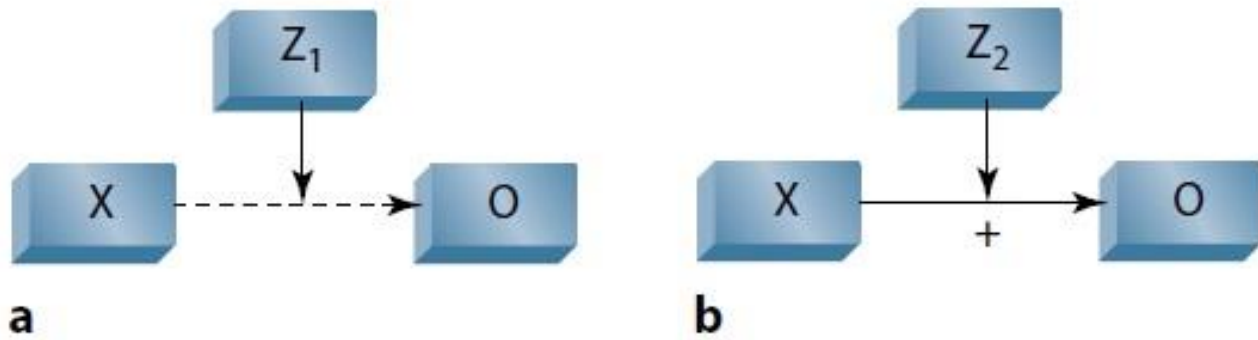


Abb. 2.10a,b Diagramm zur Veranschaulichung des Effekts von Variable X auf Variable O, der über Variable Z vermittelt wird. a Wenn Z nicht berücksichtigt wird, hat X einen signifikanten direkten Effekt auf O. b Wenn Z berücksichtigt wird, ist der direkte Effekt von X auf O nicht signifikant. Jedoch hat X einen indirekten Effekt auf O, und zwar über Z

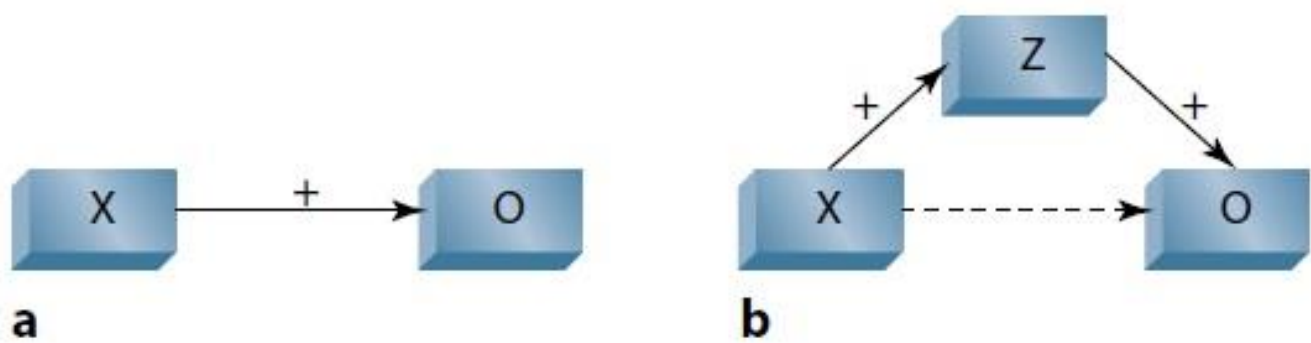


Abb. 2.11 Dieses Symbol kennzeichnet wissenschaftliche Veröffentlichungen mit präregistrierten Studien.
(Blohowiak et al. 2020. CC BY 4.0)



Abb. 2.12 Verteilung der Besucher und Besucherinnen eines Strandes in der Nähe von Durban (Südafrika) nach ethnischer Zugehörigkeit um 12:30 Uhr am 25. Dezember 1999. Anm.: Jeder rote Punkt entspricht einer Schwarzen Person; ein blauer Punkt entspricht einer Weißen Person, und ein gelber Punkt steht für eine Person mit indischer Abstammung. (Aus: Dixon & Durrheim, 2003. © 2003 The British Psychological Society, mit freundlicher Genehmigung von John Wiley & Sons)



Abb. 2.13 Implizite Messinstrumente stützen sich meist auf Computertechnologie (© Polka Dot RF / Getty Images)

