

2 Quantitative Erhebungsmethoden

Gerald Echterhoff

- 2.1 Besonderheiten psychologischer Erhebungen – 56**
 - 2.1.1 Erhellung nicht direkt beobachtbarer Phänomene – 56
 - 2.1.2 Reaktivität – 57
 - 2.1.3 Maßnahmen zur Reduzierung von Reaktivität – 58
- 2.2 Beobachten, Zählen und Messen – 62**
 - 2.2.1 Beobachten – 62
 - 2.2.2 Zählen und Messen – 64
- 2.3 Selbstberichtsverfahren: Befragung und Rating – 71**
 - 2.3.1 Befragung: Varianten und Prinzipien der Konstruktion – 74
 - 2.3.2 Ratings: Beurteilungen auf Skalen – 77
- 2.4 Testen – 80**
 - 2.4.1 Arten von Tests – 82
 - 2.4.2 Testtheorie – 84
- 2.4.3 Wünschenswerte Eigenschaften von Tests und Testitems – 86
- 2.4.4 Verfälschungen und Gegenmaßnahmen – 88
- 2.5 Biopsychologische und neurowissenschaftliche Messungen – 90**
 - 2.5.1 Messungen von Indikatoren außerhalb des zentralen Nervensystems – 93
 - 2.5.2 Messungen von Indikatoren der Aktivität des zentralen Nervensystems – 96
 - 2.5.3 Artefakte bei biopsychologischen Messungen – 103
- 2.6 Datenerhebung im Internet – 104**
 - 2.6.1 Möglichkeiten und Vorteile der Datenerhebung im Internet – 105
 - 2.6.2 Risiken und Nachteile der Datenerhebung im Internet – 108
 - 2.6.3 Hinweise zur Forschung im Internet – 109
 - 2.6.4 Experimente im WWW – 111
- 2.7 Literatur – 112**

Erhebungsmethoden der psychologischen Forschung lassen sich unterteilen in Verfahren des Beobachtens, Zählens und Messens (► Abschn. 2.2), Verfahren des Selbstberichts (► Abschn. 2.3), psychologische Tests (► Abschn. 2.4) sowie biopsychologische bzw. neurowissenschaftliche Messungen (► Abschn. 2.5). Diese Verfahren sind nicht völlig distinkt voneinander, sondern weisen Überschneidungen auf. Beispielsweise spielen Überlegungen zur Qualität einer Messung bei Selbstberichtsverfahren, Tests oder biopsychologischen Verfahren eine Rolle. Bevor wir diese Klassen von Verfahren vorstellen, gehen wir zunächst auf die Besonderheiten psychologischer Datenerhebungen ein, also auf die besonderen Bedingungen, die für psychologische Erhebungen im Unterschied zu Datenerhebungen in anderen Wissenschaften gelten (► Abschn. 2.1). Am Ende dieses Abschnitts behandeln wir zudem eine neuere und zunehmend wichtiger werdende Quelle von Daten für psychologische Untersuchungen: das Internet (► Abschn. 2.6). Hiermit ist keine eigenständige Methode angesprochen, sondern eine vergleichsweise neuartige Möglichkeit, Datenerhebungen mithilfe aller vorgestellten Methoden – mit Ausnahme der neurowissenschaftlichen Messungen – durch die Verwendung eines neuen Mediums zu realisieren. Im Zentrum des kurzen Abschnitts steht, was bei der Nutzung des Internets als Quelle von psychologischen Daten zu beachten ist, um diese auf eine methodisch möglichst korrekte Art zu gewinnen.

Um eine wissenschaftliche Fragestellung in der Psychologie in der gewünschten Breite und Tiefe zu beantworten, reicht es oft nicht aus, sich lediglich einer der hier vorgestellten Arten der Datenerhebung zu bedienen. Daher ist prinzipiell anzustreben, eine Fragestellung mit verschiedenen Datenerhebungsverfahren bzw. auf Basis verschiedener Datenquellen zu untersuchen. Jedoch ist es häufig nicht möglich oder sinn-

Zunächst geben wir einen kurzen Gesamtüberblick über das Kapitel, das Abschnitte zu den Besonderheiten psychologischer Datenerhebungen, zu den Methoden des Beobachtens, Zählens und Messens, zu Selbstberichtsverfahren, zu Tests, zu biopsychologischen Erhebungsmethoden und zur Datenerhebung im Internet enthält.

Eine Kombination verschiedener Datenerhebungsverfahren bei der Untersuchung einer Fragestellung ist grundsätzlich wünschenswert, aber nicht immer realisierbar oder aus methodischen bzw. inhaltlichen Gründen sinnvoll.

voll, innerhalb einer einzigen Studie oder Untersuchung verschiedene Methoden einzusetzen – sei es aus ökonomischen oder versuchsplanerischen Gründen. Während aus einer übergeordneten wissenschaftlichen Perspektive eine Kombination der verschiedenen Methoden grundsätzlich wünschenswert ist, werden sich Psychologen und Psychologinnen, die eine konkrete einzelne Untersuchung durchführen, oft auf wenige Methoden konzentrieren und diese gemäß der zu beantwortenden Fragestellung gezielt auswählen.

2.1 Besonderheiten psychologischer Erhebungen

Lernziele

- Den Bezug zwischen den Zielen der Psychologie und den Methoden der Datenerhebung verstehen.
- Die potenziellen Nachteile von Selbstauskünften von Versuchspersonen als Datenquelle kennenlernen.
- Das Problem der Reaktivität erkennen und anhand von Beispielen benennen.
- Verschiedene Maßnahmen zur Verringerung des Reaktivitätsproblems kennenlernen und dabei Unterschiede zwischen Feld- und Laborstudien verstehen.
- Den möglichen Beitrag indirekter bzw. impliziter Messverfahren zur Behandlung des Reaktivitätsproblems identifizieren.

In diesem Abschnitt werden Verfahren vorgestellt, mit deren Hilfe psychologische Phänomene quantitativ erfasst werden.

Ein zentrales Ziel psychologischer Forschung ist die Erhellung der **Black Box** (des nicht direkt beobachtbaren psychischen Innenlebens).

Beobachtbares Verhalten ist psychologisch bedeutsam, wenn es im Hinblick auf zugrunde liegende psychische Phänomene interpretiert werden kann.

Bei der Datenerhebung im Bereich quantitativer Methoden geht es stets darum, das Erleben und Verhalten von Menschen, z. B. ihre Kognitionen, Gefühle, Motivationen und ihren Umgang miteinander, möglichst genau zu erfassen, und zwar in numerischer Form. Die Vor- und Nachteile quantitativer Informationen haben wir eingangs diskutiert. In diesem Abschnitt stellen wir Verfahren vor, mit deren Hilfe in der Psychologie zu solchen quantitativen Informationen gelangt werden kann.

2.1.1 Erhellung nicht direkt beobachtbarer Phänomene

Besonderheiten psychologischer Datenerhebung wurden schon zu Anfang dieses Buchs (► Kap. 1) angesprochen: Der psychologischen Forschung geht es im Kern um die Erhellung nicht direkt beobachtbarer Phänomene, vor allem der Vorgänge »im Kopf« von Menschen. Die wissenschaftliche Psychologie versucht spätestens seit dem Ende der Dominanz behavioristischer Ansätze (in den 1950er/1960er Jahren), in die sog. »Black Box«, also in das Innere des psychischen Systems vorzudringen und die darin ablaufenden Prozesse zu erhellen. Was erlaubt es Menschen, sich an ihre Kindheit, Lernstoff aus dem Geschichtsunterricht oder an Gedichte zu erinnern? Wie kommen Menschen dazu, andere Personen als mehr oder weniger attraktiv einzuschätzen? Was passiert beim Lösen von Problemen, also etwa bei der Planung der besten Reiseverbindung von X nach Y oder der Suche nach bestimmten Informationen im Internet?

Viele der Prozesse, die aus psychologischer Sicht bei all diesen Beispielen relevant sind, finden innerhalb des psychischen Systems der sich erinnernden, urteilenden und denkenden Individuen statt. Dass sich solche Vorgänge in einem von außen nicht direkt zugänglichen Inneren abspielen, ist besonders augenfällig an der Herkunft des Wortes »erinnern« ablesbar, das sich vom althochdeutschen Raumadjektiv »innaro« (das/der Innere, innerer) herleitet. Es sei angemerkt, dass auch schon direkt beobachtbares Verhalten von Menschen informativ sein kann. Jedoch interessiert es die heutige Psychologie primär im Hinblick darauf, welche Schlüsse daraus auf die zugrunde liegenden Strukturen, Prozesse und Mechanismen des psychischen Geschehens gezogen werden können. Offenes Verhalten bleibt aus psychologischer Sicht bedeutungslos, solange es nicht im Hinblick auf das Innenleben der menschlichen Psyche interpretiert wird.

2.1 · Besonderheiten psychologischer Erhebungen

Zu Beginn von Untersuchungen des Innenlebens der Black Box steht die Frage, ob die interessierenden psychischen Prozesse den untersuchten Personen selbst zugänglich und bekannt sind und ob die Personen über diese Prozesse auch ohne unerwünschte Verzerrungen Auskunft geben können, diese also den Forschenden in angemessener Form kommunizieren können. Kritisch ist stets zu überprüfen, ob solche **Selbstauskünfte** von Personen hinreichend reliabel und valide sind (► Abschn. 1.4). Falls diese Fragen – zumindest tendenziell – bejaht werden können, so bietet sich eine Datenerhebung mithilfe von Selbstberichtsverfahren an (► Abschn. 2.3). Bei solchen Verfahren, zu denen etwa eine mündliche Befragung oder die Selbstbeurteilung auf Ratingskalen gehören, berichten die Untersuchungsteilnehmer und -teilnehmerinnen über ihr eigenes psychisches Geschehen, über innere Vorgänge, beispielsweise über ihre Gefühle, Denkprozesse, Motive, Erinnerungen etc. Falls Auskünfte der untersuchten Personen als nicht hinreichend reliabel und valide Quellen für Informationen über die interessierenden psychischen Prozesse gelten können (s. unten), so sind Selbstberichtsverfahren zu vermeiden und die relevanten Daten stattdessen mithilfe von Verfahren zu erheben, die ohne Selbstauskünfte der Befragten auskommen (z. B. Beobachtungs- und Messverfahren oder biopsychologische Methoden; ► Abschn. 2.2 und ► Abschn. 2.5).

2.1.2 Reaktivität

Eine zentrale Überlegung zu Beginn der Planung einer psychologischen Studie betrifft die mögliche **Reaktivität** des Untersuchungsgegenstands (■ Abb. 2.1).

Definition

Reaktivität bei psychologischen Datenerhebungen bedeutet die Veränderung bzw. Verzerrung der erhobenen Daten schon aufgrund der Kenntnis der untersuchten Personen darüber, dass sie Gegenstand einer Untersuchung sind. Aufgrund von Reaktivität verändert sich die Beschaffenheit des zu erforschenden Gegenstands.



Wenn eine Person beim Lösen eines komplexen Problems im Alltag befragt wird, was ihr gerade durch den Kopf geht oder wie sie das Problem anzugehen versucht, kann der interessierende psychische Prozess genau durch diesen Akt der Befragung verändert werden. Ähnliches kann in sozialpsychologischen Studien zur zwischenmenschlichen Anziehung geschehen, wenn etwa zwei untersuchte Personen während des Kennenlernens befragt werden, wie sympathisch oder attraktiv sie einander finden. Wie wir sicherlich aus eigener Erfahrung wissen, kann sich unser Erleben grundlegend verändern, wenn wir dazu befragt werden oder wissen, dass wir beobachtet werden (► Beispiel).

Bei der Forschungsplanung ist zu berücksichtigen, ob die untersuchten Personen Zugang zu den relevanten psychischen Prozessen haben und ob sie über diese Prozesse ohne unerwünschte Verzerrungen selbst Auskunft geben können; nur dann sind Selbstauskünfte prinzipiell reliabel und valide Datenquellen.

Von Beginn an ist die Reaktivität des Untersuchungsgegenstands zu beobachten.

► Definition Reaktivität

■ **Abb. 2.1** Reaktivitätsproblem – Schnäppchenjäger

Datenerhebungen in Psychologie und Sozialwissenschaften verändern oft schon den Gegenstand; hierfür gibt es viele Beispiele.

Beispiel**Reaktivität: Aufforderungscharakteristika beim Hawthorne-Effekt**

Ein populäres Beispiel für Reaktivität ist der Hawthorne-Effekt. In einer klassischen Studie von Roethlisberger und Dickson (1939) in den Hawthorne-Werken der Western Electric Company in den USA zeigte sich, dass schon die bloße wissenschaftliche Untersuchung der Arbeiterinnen zu einer Steigerung der Produktivität führte. Diese Produktivitätssteigerung war unabhängig davon, welche Veränderung der Arbeitsbedingungen von den Forschern imple-

mentiert wurde: Die Arbeiterinnen zeigten z. B. eine verbesserte Leistung nicht nur bei einer Erhöhung, sondern auch bei einer Verringerung der Beleuchtungsstärke am Arbeitsplatz! Es handelt sich hierbei um einen Effekt, der in der psychologischen Literatur auf die Bewertungserwartung der untersuchten Personen oder Aufforderungscharakteristika (»demand characteristics«) der Untersuchung zurückgeführt wird.

Die **Reaktivität** ist eine beständige Herausforderung an die psychologische Forschung, da das Bewusstsein, beobachtet zu werden, das Verhalten der Beobachteten beeinflussen kann.

Das Problem der Reaktivität stellt sich in allen Forschungsbereichen, in denen Menschen der Untersuchungsgegenstand sind. Viele Befunde zeigen, dass das bloße Bewusstsein, beobachtet oder wissenschaftlich untersucht zu werden, das Verhalten und Erleben der untersuchten Personen beeinflussen kann. Wenn Menschen annehmen, dass sie von anderen beobachtet und bewertet werden, versuchen sie beispielsweise, durch ihr Verhalten beim Beobachter einen positiven Eindruck zu hinterlassen (► Beispiel).

Beispiel**Reaktivität: Aufforderungscharakteristika beim Waffeneffekt**

Auch experimentelle Laboruntersuchungen können Aufforderungscharakteristika beinhalten, d. h. mehr oder weniger subtile Hinweise darauf, wie sich die Versuchspersonen verhalten sollten. In den bekannten Experimenten zum Waffeneffekt (Berkowitz & LePage, 1967) hatten die (männlichen) Versuchspersonen die Option, einen Mitarbeiter der Versuchsleitung, der den Versuchspersonen jedoch als ein regulärer Versuchsteilnehmer vorgestellt wird, mit einem elektrischen Schock zu bestrafen. Dieser Konfident der Versuchsleitung hatte ihnen zuvor ebenfalls einen Schock erteilt. Die Versuchspersonen wählten häufiger die Option des Schocks, wenn sich im Versuchsraum ein Revolver und ein Gewehr (versus ein Paar Badmintonschläger) befanden. Diesen Befund interpretieren Berkowitz und LePage als Beleg für die Hypothese, dass Hinweisreize, die mit Aggression assoziiert sind, aggressive Gedanken auslösen können, die zu

aggressivem Verhalten führen. Jedoch machten Kritiker dieser Interpretation geltend, dass der festgestellte Unterschied alleine durch die Aufforderungscharakteristika des Versuchsaufbaus erklärbar ist und somit nicht auf aggressive Gedanken der Versuchspersonen zurückgeführt werden muss. Dieser Alternativerklärung zufolge signalisieren die Waffen im Versuchsraum den Versuchspersonen, wie sie sich in der Situation verhalten sollen, nämlich aggressiv, indem sie der anderen Person einen Elektroschock erteilen. Es sei angemerkt, dass Ergebnisse nachfolgender Studien gegen die Erklärung durch Aufforderungscharakteristika sprechen. So zeigte eine Studie von Turner, Simons, Berkowitz und Frodi (1977), dass der Waffeneffekt reduziert wurde, wenn den Versuchspersonen die Einsicht in die zugrunde liegende Hypothese erleichtert wurde.

Dem Problem der Reaktivität ist aufmerksam und mithilfe geeigneter Strategien zu begegnen.

2.1.3 Maßnahmen zur Reduzierung von Reaktivität

Die Probleme, die sich aus der Reaktivität der menschlichen Untersuchungssubjekte ergeben, sind seit Langem bekannt. Psychologinnen und Psychologen müssen sich stets darüber im Klaren sein, ob und in welcher Hinsicht ihre Datenerhebung reaktiv sein könnte und welche nichtreaktiven Verfahren alternativ zur Verfügung stehen (zu solchen Verfahren bei qualitativen Beobachtungen vgl. ► Abschn. 6.2; für eine Übersicht über nichtreaktive Verfahren vgl. Webb, Campbell, Schwartz, Sechrest & Grove, 1981). Im Folgenden nennen wir einige der wichtigsten Strategien, derer sich in der Forschung bedient werden kann, um das Problem der Reaktivität zu verringern (■ Tab. 2.1).

■ **Tab. 2.1** Maßnahmen zur Verringerung des Reaktivitätsproblems

Maßnahme	Erläuterung
Untersuchte in Unkenntnis darüber lassen, dass sie untersucht werden	Nur in Feld-, Archiv- oder Internetstudien praktikabel, nicht in Laborstudien; kann ethisch problematisch sein
Untersuchten Anonymität zusichern	Besonders wichtig bei der Erhebung von persönlichen Meinungen, Einstellungen oder anderen sensiblen Daten
Untersuchten eine Coverstory über den Untersuchungszweck mitteilen	Wichtig in hypothesenprüfenden Studien, in denen die Untersuchten die erforschten Verhaltensweisen kontrollieren oder gezielt steuern können; dies ist aber jeweils ethisch zu reflektieren (► Abschn. 1.7.2)
Maße einsetzen, die die Untersuchten nicht kontrollieren oder beeinflussen können (nichtreaktive Messverfahren)	Wird angenommen für biopsychologische Maße, die willkürlich nicht oder kaum steuerbare physiologische Vorgänge erfassen, wie etwa die Messung des Kortisolspiegels oder bildgebende Verfahren zur Registrierung der Gehirnaktivität
Indirekte/implizite Messverfahren einsetzen	Neuerer Ansatz, bei dem die Untersuchten aus der Art der gemessenen Verhaltensweisen (oft Reaktionszeiten) nicht oder nur schwer auf das untersuchte psychologische Konstrukt (z. B. Vorurteile gegenüber Fremdgruppen) schließen können; die Verfahren (z. B. der IAT) werden noch kritisch diskutiert, sind aber eine interessante Option

Maßnahme 1: Unkenntnis der Untersuchten darüber, dass sie Gegenstand einer Untersuchung sind Die effektivste Möglichkeit zur Behandlung des Reaktivitätsproblems besteht darin, die Versuchspersonen gänzlich darüber im Unklaren zu lassen, dass sie untersucht werden. Diese Möglichkeit besteht bei Studien mit archivarischem Material (z. B. Tagebüchern), Feldstudien oder Studien im Internet (WWW), jedoch nicht bei Laboruntersuchungen. Denn bei Laborstudien ist es natürlich unvermeidlich, die Teilnehmenden darüber in Kenntnis zu setzen, dass sie das Labor zum Zweck der Durchführung einer Studie aufsuchen. Studien im Feld, also in natürlicher Umgebung, können hingegen stattfinden, ohne dass die Untersuchten informiert werden, dass sie Gegenstand einer Untersuchung sind. In der Tat liegt der Reiz und Vorteil von Feldstudien gerade darin, Erleben und Verhalten von Menschen ohne Eingriff durch den Forscher in natürlichen Umgebungen zu erheben (► Beispiel).

Die Untersuchten in **Unkenntnis** darüber zu lassen, dass sie Gegenstand einer Untersuchung sind, ist die effektivste Maßnahme, jedoch bei Laborstudien nicht realisierbar.

Beispiel

Unkenntnis der Untersuchten über stattfindende Untersuchung in einer Feldstudie

In einem sozialpsychologischen Feldexperiment prüften Ellsworth, Carlsmith und Henson (1972) beispielsweise die Hypothese, dass ein relativ langer Blickkontakt Menschen unangenehm ist und zu Fluchtverhalten führt. In einer Versuchsbedingung blickte ein an der Kreuzung postierter Mitarbeiter (der als solcher für die Autofahrer nicht erkennbar

war) intensiv in ihre Richtung. In einer Kontrollbedingung schaute der Mitarbeiter von ihnen weg, also in eine andere Richtung. Ellsworth et al. stellten fest, dass Autofahrer nach dem Wechsel einer Ampel auf Grün in der ersten Bedingung schneller wegfuhr.

Ob nichtreaktive Feldstudien wie die von Ellsworth et al. (1972; ► Beispiel) durchgeführt werden sollten, hängt u. a. von forschungsethischen Überlegungen (► Abschn. 1.7) ab. Ein Vorgehen wäre etwa nicht vertretbar, wenn den Untersuchten durch die Untersuchung ernstzunehmende Nachteile entstehen können (z. B. wenn in der genannten Studie von Ellsworth et al. befürchtet worden wäre, dass der Blickkontakt zu einem deutlich höheren Unfallrisiko führen könnte).

Die Durchführbarkeit nichtreaktiver Feldstudien hängt von forschungsethischen Rahmenbedingungen ab.

Maßnahme 2: Zusicherung von Anonymität Da in Laborstudien keine völlige Unkenntnis der Teilnehmenden über die Durchführung einer Studie realisierbar ist, müssen ggf. andere Wege zur Verringerung des Reaktivitätsproblems beschritten werden. Eine simple Maßnahme besteht darin, den Versuchsteilnehmerinnen und -teilnehmern

Die Zusicherung von **Anonymität** ist vor allem in Laborstudien und Umfragen von Bedeutung.

Die Formulierung einer **Coverstory** ist bei hypothesenprüfenden Untersuchungen sinnvoll, falls die Kenntnis der Hypothesen durch die Untersuchten die Validität oder Präzision der Hypothesenprüfung gefährdet.

Anonymität zuzusichern und die hohe Bedeutung unverfälschter Aussagen und Verhaltensweisen deutlich zu machen. Solche Maßnahmen sind besonders bei Umfragen unerlässlich, bei denen persönliche Meinungen, Einstellungen oder andere sensible Daten erfragt werden und die Forschenden auf die Ehrlichkeit und die Kooperation der Befragten angewiesen sind.

Maßnahme 3: Formulierung einer Coverstory In Untersuchungen, die der Prüfung von Hypothesen dienen, kann zudem eine sog. **Coverstory** eingesetzt werden, die gezielt bestimmte Erwartungen über die Fragestellung und den Versuchshintergrund bei den Versuchspersonen induziert (► Beispiel). Wenn die Kenntnis der Hypothesen die Validität oder Präzision der Untersuchung (► Abschn. 3.2.5) gefährdet, dann sollte die Coverstory einen anderen als den wirklichen Hintergrund beinhalten, die Versuchspersonen also über den Sinn und Zweck des Experiments täuschen (zur ethischen Problematik ► Abschn. 1.7).

Beispiel

Coverstory in sozialpsychologischen Studien

In Milgrams (1963) klassischer Studie zu Gehorsam besagte die Coverstory z. B., dass die Versuchspersonen die Rolle eines Lehrers übernehmen sollten, dessen Aufgabe es sei, einer lernenden Person Wortpaare beizubringen. In den Studien zum Waffeneffekt von Berkowitz und LePage (1967) wurde den Versuchsteilnehmern mitgeteilt, es ging um die

Bewertung von Problemlösevorschlügen durch milde elektrische Schocks und die Reaktion der bewerteten Person auf solche Schocks. Somit sollten die Teilnehmer von der wahren Fragestellung abgelenkt werden, die sich nicht auf die Reaktionen der bewerteten Person, sondern auf das (aggressive) Verhalten der bewertenden Person richtete.

Durch die Verwendung **nichtreaktiver Messverfahren** können die Ergebnisse von den Untersuchten nicht bzw. kaum kontrolliert oder beeinflusst werden.

Maßnahme 4: Verwendung nichtreaktiver Messverfahren Eine weitere Möglichkeit, das Problem der Reaktivität in Laborstudien zumindest zu verringern, besteht im Einsatz von Maßen oder Datenerhebungsmethoden, deren Befunde eine untersuchte Person nicht kontrollieren oder beeinflussen kann. Bei einigen Verfahren können die Untersuchten vermutlich keinen oder keinen nennenswerten Einfluss auf das Ergebnis der Messung nehmen, selbst wenn sie den genauen Zweck der Messung kennen. Hierzu zählen biopsychologische Maße, die willkürlich nicht oder kaum steuerbare physiologische Vorgänge erfassen, wie etwa die Registrierung der elektrodermalen Aktivität oder des Kortisolspiegels sowie bildgebende Verfahren zur Registrierung der Gehirnaktivität (► Abschn. 2.5).

Durch die Verwendung **indirekter/impliziter Messverfahren** können Verhaltensaspekte (z. B. Reaktionszeiten) ausgewertet werden, deren Zusammenhang mit der Fragestellung den untersuchten Personen in der Regel verborgen ist.

Maßnahme 5: Verwendung indirekter/impliziter Messverfahren Zudem werden in jüngerer Zeit Maße entwickelt, die das Reaktivitätsproblem reduzieren, indem den untersuchten Personen ein anderer als der wahre Zweck der Messung suggeriert wird. Diese Maße werden als **indirekte** bzw. **implizite** Messverfahren bezeichnet. Bei diesen Maßen werden Verhaltensaspekte ausgewertet, die die untersuchten Personen vermutlich nicht mit der wirklichen Fragestellung oder den Forschungshypothesen in Zusammenhang bringen können. Bei einigen neueren Verfahren wird etwa aus der Antwortzeit bzw. den Reaktionslatenzen von Versuchspersonen bei der Beantwortung von Fragen oder Aufgaben auf die zugrunde liegenden kognitiven Strukturen und Prozesse geschlossen (► Beispiel). Da diese Verfahren nicht auf einer expliziten (offenen) Befragung der Versuchspersonen nach ihren psychischen Zuständen basieren, werden sie als implizit bezeichnet.

Beispiel**Beispiel für ein implizites Verfahren zur Einstellungsmessung: Der IAT**

Das bekannteste unter den impliziten Verfahren ist der »Implicit Association Test« (IAT; Greenwald, McGhee & Schwartz, 1998). Mithilfe des IAT soll erfasst werden, wie schnell Personen verschiedene Stimuli miteinander assoziieren. Der IAT basiert auf der Grundidee, dass Menschen um so schneller auf Stimuli reagieren können, je stärker zu dem Zeitpunkt bereits verwandte kognitive Konzepte und Strukturen »in ihrem Kopf« mitaktiviert oder voraktiviert sind. Wenn eine deutsche Versuchsperson beispielsweise schneller mit derselben Reaktion (z. B. Drücken ein und derselben Taste) auf türkische Namen und negative Begriffe wie »Krieg« oder »Ärger« antwortet als auf deutsche Namen gepaart mit negativen Begriffen, dann wird dies als Indikator dafür interpretiert, dass die Person eher Negatives mit der Kategorie

»türkisch« verbindet. Dieses Reaktionszeitmuster zeigt der Testlogik zufolge eine implizit negative Einstellung gegenüber Türken an. Viele Versuchspersonen gehen nicht davon aus, dass durch ihre Reaktionsgeschwindigkeit auf diese Assoziationen geschlossen wird, und können zudem die Schnelligkeit einer spontanen Reaktion kaum kontrollieren. Daher wird in der Forschung angenommen, dass dieses Messverfahren nicht durch Erwartungen oder eine bewusste Antwortkontrolle der Versuchspersonen verfälscht werden kann. Zwar wäre eine gezielte Beeinflussung von Reaktionen durch die Untersuchten möglich; jedoch kann diese ausgeschlossen werden, wenn die Untersuchten die Reaktionen nicht mit dem eigentlichen Ziel der Messung in Verbindung bringen.

Obwohl die Interpretation und Güte solcher impliziter Erhebungsmethoden, vor allem des IAT (► Beispiel), nicht unumstritten sind (z. B. Blanton & Jaccard, 2006), so haben sie die wissenschaftliche Diskussion über nichtreaktive Zugänge zum psychischen Geschehen auf jeden Fall erheblich vorangebracht und der Forschung stimulierende Impulse verliehen. Da ein abschließendes Urteil über den Nutzen impliziter Messverfahren derzeit nicht gefällt werden kann, ist es sicherlich lohnend, diese bei der Prüfung nichtreaktiver Erhebungsalternativen auch in Erwägung zu ziehen.

Eine endgültige Bewertung impliziter Messverfahren ist noch nicht möglich. Sie haben die Forschung stimuliert, werden aber auch weiterhin kontrovers diskutiert.

? Kontrollfragen

1. In welchen Hinsichten unterscheiden sich Datenerhebungen in der Psychologie von Datenerhebungen in anderen empirischen Wissenschaften? Welche besonderen Herausforderungen ergeben sich für psychologische Messungen?
2. Was ist bei der Erhebung von Selbstauskünften von Versuchspersonen zu beachten?
3. Worin besteht das Problem der Reaktivität und wie kann es verringert werden?
4. Inwiefern unterscheiden sich die Möglichkeiten zur Behandlung des Reaktivitätsproblems zwischen Feld- und Laborstudien?
5. Wozu dient die Formulierung einer Cover Story in psychologischen Untersuchungen?
6. Welche Rolle können Reaktionszeiten bei der Behandlung des Reaktivitätsproblems in psychologischen Datenerhebungen spielen? Stellen Sie einen Bezug zu neueren impliziten Messverfahren her.

Pelham, B., & Blanton, H. (2007). *Conducting research in psychology: Measuring the weight of smoke* (3rd ed.). Belmont: Thomson Wadsworth.

Webb, E. J., Campbell, D. T., Schwartz, R. F., Sechrest, L. & Grove, J. B. (1981). *Nonreactive measures in the social sciences*. Boston: Houghton Mifflin.

► Weiterführende Literatur

2.2 Beobachten, Zählen und Messen

Lernziele

- Beobachtung im Alltag von wissenschaftlicher Beobachtung unterscheiden.
- Verschiedene Arten von Beobachtungsstichproben kennenlernen.
- Die grundlegenden Verfahren des Zählens und Messens definieren und unterscheiden.
- Grundlagen der Messtheorie kennenlernen.
- Verschiedene Arten von Merkmalen (diskrete, stetige) unterscheiden und den Bezug dieser Unterscheidung zur Wahl der Datenerhebungsverfahren würdigen.
- Die unterschiedlichen Skalenniveaus verstehen.
- Verstehen, was zulässige Transformationen sind und weshalb diese für psychologische Datenerhebungen bedeutsam sind.

Die **wissenschaftliche Beobachtung**, d. h. die systematische und regelgeleitete Registrierung des Auftretens bzw. der Ausprägung von ausgewählten, psychologisch relevanten Merkmalen oder Ereignissen, ist – wie die visuelle Wahrnehmung überhaupt – selektiv und konstruktiv.

Bei der wissenschaftlichen Beobachtung wird, im Unterschied zur Beobachtung im Alltag, ihre Selektivität und Konstruktivität bedacht und möglichst kontrolliert.

► Definition Wissenschaftliche Beobachtung

Im Anschluss an die einführenden Bemerkungen zu den Besonderheiten psychologischer Datenerhebungen und insbesondere dem Problem der Reaktivität folgt nun die Darstellung der gängigen Erhebungsmethoden.

2.2.1 Beobachten

Selektivität und Konstruktivität

Jede Datenerhebung beinhaltet den Vorgang der visuellen Wahrnehmung, also Beobachtung. Mit **Beobachtung** als eigener Datenerhebungsmethode ist jedoch etwas weniger Allgemeines gemeint, nämlich die systematische und regelgeleitete Registrierung des Auftretens bzw. der Ausprägung von ausgewählten, psychologisch relevanten Merkmalen oder Ereignissen (vor allem von Verhaltensweisen von Menschen). Die Beobachtung ist mit den typischen Herausforderungen der visuellen Wahrnehmung konfrontiert, also mit der Selektion von Informationen aus einem komplexen, sich ständig verändernden Strom von permanent eintreffenden visuellen Daten und deren Organisation zu bedeutungsvollen und stabilen Einheiten. Viele Ansätze und Befunde der Wahrnehmungspsychologie weisen darauf hin, dass diese Prozesse in einem beachtlichen Maß konstruktiv und aktiv gestaltend auftreten (z. B. Neisser, 1979). Jede Beobachterin und jeder Beobachter erzeugt also ein Bild der Wirklichkeit, das u. a. von ihren bzw. seinen Einstellungen, Erwartungen und Vorerfahrungen abhängt. Solche Beobachtereinflüsse sind bei der Wahrnehmung anderer Menschen, also komplexer und zugleich höchst relevanter Objekte, besonders stark ausgeprägt (z. B. Higgins & Bargh, 1987).

In der Psychologie ist aus den genannten Gründen in einem erheblichen Ausmaß mit beobachterabhängigen Erwartungs- und Einstellungseffekten zu rechnen. Beobachtung als psychologische Datenerhebungsmethode unterscheidet sich von der (ununterbrochen stattfindenden) Beobachtung im Alltag darin, dass im ersteren Fall die beobachterabhängige Selektivität und Konstruktivität berücksichtigt und möglichst kontrolliert werden soll. Um dies zu leisten, erfolgt die wissenschaftliche Beobachtung auf systematische und regelgeleitete Weise (Bortz & Döring, 2006).

Definition

Wissenschaftliche Beobachtung ist die systematische und regelgeleitete Registrierung des Auftretens bzw. der Ausprägung von ausgewählten, psychologisch relevanten Merkmalen oder Ereignissen. Sie folgt einem zuvor festgelegten Beobachtungsplan, der festlegt,

1. was beobachtet werden soll (Kategorien für das/die interessierende/n Ereignis/se oder Merkmal/e);



2. welche Aspekte weniger oder nicht relevant sind;
3. welchen Interpretationsspielraum der Beobachtende bei der Beobachtung hat;
4. wann, wie lange und wo die Beobachtung erfolgt (Zeitpunkte, Zeiträume, Situationen);
5. auf welche Weise das Beobachtete registriert und protokolliert wird.

Werden mehrere Personen zur Beobachtung eingesetzt, ist zu entscheiden, ob und wie die Beobachtungsaufgabe unter ihnen aufgeteilt wird und auf welche Aspekte sie jeweils achten sollen. Das Ziel dieser Festlegungen besteht darin, die Reliabilität, Validität und Objektivität (► Abschn. 1.4) der durch die Beobachtung erhobenen Daten zu sichern. Ein Beispiel für ein Beobachtungssystem zur Untersuchung der Interaktion von Schülern und Lehrern ist in ■ Tab. 2.2 wiedergegeben.

Auswahl von Stichproben

Auch wenn diese Gütekriterien gesichert sind, so kann Beobachtung, auch wenn sie sich technischer Hilfsmittel und Apparate wie etwa einer Videoaufzeichnung bedient, nie die Gesamtheit aller Ereignisse erfassen und muss sich daher mit mehr oder weniger großen Ausschnitten begnügen. Die Beobachtung muss also auf Stichproben des Gesamtgeschehens beschränkt bleiben; damit stellt sich wie auch bei der Auswahl von Personenstichproben die Frage der Repräsentativität (► Abschn. 3.1.3). Die stichprobenartig ausgewählten Beobachtungseinheiten sollten natürlich hinreichend typisch und repräsentativ sein.

Zwei Hauptformen der **Stichprobenziehung** sind zu unterscheiden: die Zeitstichprobe und die Ereignisstichprobe. Die beiden Arten der Stichprobenziehung unterscheiden sich darin, an welchen Punkten der zu untersuchende Gegenstand (z. B. Verhaltensweisen von Angestellten, die unter Zeitdruck stehen) beobachtet wird.

■ **Tab. 2.2** Beispiel für ein Beobachtungssystem: Das »Verbal Interaction Category System« zur Beobachtung der Lehrer-Schüler-Interaktion von Amidon und Hunter (1967; dtsh. Grell, 1980)

Kategorie	Unterkategorie (Verhaltensaspekte)
Vom Lehrer ausgehendes Verbalverhalten	Äußert Informationen oder Meinungen
	Gibt Anweisungen
	Stellt enge Fragen
	Stellt weite Fragen
Lehrererwiderung	Akzeptiert a) Gedanken, b) Verhalten, c) Gefühle
	Lehnt a) Gedanken, b) Verhalten c) Gefühle ab
Schülererwiderung	Voraussagbare Antwort an Lehrer
	Nichtvoraussagbare Antwort an Lehrer
	Antwort an Mitschüler
Vom Schüler ausgehendes Verbalverhalten	Beginnt Gespräch mit Lehrer
	Beginnt Gespräch mit Mitschüler
Anderes	Stille
	Durcheinander

Ein Beobachtungssystem soll die Reliabilität, Validität und Objektivität der Daten sichern.

Da Beobachtung niemals alles erfassen kann, müssen **Stichproben** des Gesamtgeschehens ausgewählt werden.

Bei der **Zeitstichprobe** werden Beobachtungen in festen Intervallen aufgezeichnet.

Bei der **Ereignisstichprobe** wird das Auftreten, die Auftretensdauer (oder -häufigkeit) von definierten Ereignissen aufgezeichnet.

Zeit- und Ereignisstichproben weisen Vor- und Nachteile auf.

Beobachtung ist auch ein fester Bestandteil qualitativer Methoden.

Die Häufigkeit beobachteter Ereignisse wird durch Zählen bestimmt.

Zählen gibt über Häufigkeiten Auskunft, während Messen eine feinere Erfassung von Merkmalsausprägungen erlaubt (d. h. wie stark ein Merkmal vorliegt).

Zeitstichprobe Bei der Zeitstichprobe werden Beobachtungen in festgelegten Zeitabständen protokolliert, z. B. in 20-Sekunden-Intervallen. So könnte etwa in festen Zeitintervallen aufgezeichnet werden, was beobachtete Angestellte unter Zeitdruck tun, z. B. Text in einen Computer eingeben, Informationen aufnehmen oder mit anderen Mitarbeitern sprechen.

Ereignisstichprobe Im Fall einer Ereignisstichprobe wird hingegen registriert, ob, wie lange oder wie oft bestimmte zuvor definierte Ereignisse auftreten, ohne dass einem festen zeitlichen Rhythmus gefolgt wird. Die Beobachterin bzw. der Beobachter wartet hierbei darauf, dass sich ein zuvor definiertes Ereignis zeigt, und protokolliert dann, um welches Ereignis es sich handelt und ggf. wie lange es andauert. Bei der Beobachtung von Angestellten unter Zeitdruck würde etwa auf das Auftreten von bestimmten interessierenden Verhaltensweisen (Texteingabe am Computer; Kommunikation mit einem Vorgesetzten oder gleichgestellten Mitarbeiter; Wechsel des Aufenthaltsorts etc.) geachtet und diese dann entsprechend protokolliert.

Im Vergleich zur Zeitstichprobe ist es bei der Ereignisstichprobe eher möglich, auch relativ selten auftretende Ereignisse zu erfassen. Zudem erlaubt die Ereignisstichprobe eher die gezielte Überprüfung von Hypothesen zum Zusammenhang (Kontingenz) von aufeinander folgenden Ereignissen: Welche Ereignisse folgen typischerweise aufeinander, welches Ereignis geht einem anderen zumeist unmittelbar voraus? Während Ereignisstichproben eher der gezielten Untersuchung ausgewählter Verhaltensweisen oder deren Kombinationen dienen, bieten Zeitstichproben also einen breiteren und umfassenderen Überblick über einen Gesamtverlauf von Ereignissen.

Beobachtung in der qualitativen Forschung

Beobachtung als Erhebungsmethode ist im Unterschied zum Messen oder Testen nicht auf den Bereich der quantitativen Methoden festgelegt. Im Gegenteil: Beobachtungsverfahren werden häufiger im Bereich qualitativer Methoden eingesetzt. Daher sei an dieser Stelle auf den entsprechenden Abschnitt zu qualitativen Datenerhebungsverfahren verwiesen (► Abschn. 6.2.1). Dort finden sich weitere Erläuterungen und Dimensionen zur Unterscheidung verschiedener Beobachtungsformen (z. B. Grad der Standardisierung; teilnehmende vs. nicht teilnehmende Beobachtung).

Um die Daten aus einer Beobachtung quantitativ zu verwerten, ist es zunächst erforderlich, die Häufigkeiten der beobachteten Ereignisse zu bestimmen. Hierzu dient die elementare Operation des Zählens. In dem Beispiel für ein Beobachtungssystem aus **Tab. 2.2** (»Verbal Interaction Category System«) würde also beispielsweise gezählt, wie häufig der Lehrer in einer bestimmten Situation oder in einem bestimmten Zeitraum Fragen stellt oder wie häufig Schüler Gespräche mit Mitschülern beginnen. Um das Zählen geht es – neben der differenzierteren Option des Messens – im folgenden Abschnitt.

2.2.2 Zählen und Messen

Elementare Schritte der quantitativen Datenerhebung, die an die Beobachtung anschließen bzw. mit ihr verbunden sind, sind das Zählen und Messen: In vielen Studien interessiert nicht nur, dass ein Ereignis (z. B. aggressives Verhalten wie etwa eine Beleidigung) aufgetreten ist, sondern *wie oft* ein Ereignis eintritt. Ist man an den Häufigkeiten von Ereignissen interessiert, ist das Auftreten der relevanten Ereignisse zu *zählen*. Darüber hinaus lassen sich Merkmale auch mithilfe von Messungen quantifizieren. Im Vergleich zum Zählen ermöglicht das Messen eine feinere, abgestufte Erfassung von Merkmalsausprägungen, also z. B. der Intensität bzw. des Ausmaßes aggressiven Verhaltens. Messen kann somit darüber Auskunft geben, wie deutlich oder wie stark ein Merkmal bzw. ein Ereignis vorliegt.



<http://www.springer.com/978-3-642-34361-2>

Forschungsmethoden in Psychologie und
Sozialwissenschaften für Bachelor

Hussy, W.; Schreier, M.; Echterhoff, G.

2013, X, 318 S. Mit Online-Extras., Softcover

ISBN: 978-3-642-34361-2