

METHODEN DER INFORMATIONSGEWINNUNG

Verhaltensbeobachtung

Bedeutung und Merkmale

Mit Hilfe der Beobachtung können Informationen erlangt werden, die sonst nicht oder nur schwer erfasst werden können. Das Messinstrument ist in allen Fällen der Beobachter, dessen Beobachtung durch seine Erfahrungen, Bedürfnisse oder Einstellungen beeinflusst wird.

KÖNIG (1967) grenzt die wissenschaftliche von der naiven Beobachtung durch die Prinzipien der **Gezieltheit**, der **Konstanz** und der **Kontrolle** ab:

- **Gezieltheit:** Beobachtung dient einem bestimmten Zweck und gibt sich nicht wechselnden Reizen hin; der wissenschaftliche Beobachter muss vorher Hypothesen aufstellen, die er durch die Beobachtung überprüfen will
- **Konstanz:** Man begnügt sich nicht mit flüchtigen Eindrücken und überbewertet nicht zufällig auftretende kurzfristige Erscheinungen; das Konstanzprinzip ist anders zu bewerten, wenn situative Einflüsse auf das Verhalten beobachtet werden sollen
- **Kontrolle:** Aussagen, die man aufgrund von Beobachtungen über bestimmte Personengruppen macht, werden auch durch Kontrollbeobachtungen an anderen Personengruppen darauf hin überprüft, ob sie spezifisch für eine Gruppe sind

Formen der wissenschaftlichen Beobachtung

a) nicht-standardisierte – standardisierte

Die **nicht-standardisierte Beobachtung** darf nicht mit der naiven verwechselt werden. Die Fragestellung ist nur weich, aber vorhanden. Einsatz meist zu Beginn einer Forschungsarbeit, wenn man sich orientieren will, welche Phänomene auftreten.

Erst danach führt man detaillierte Beobachtungsschemata ein, die die zu beobachtenden Sachverhalte, Verhaltensmuster und Handlungsdimensionen sowie die Art ihrer Registrierung in der Form der **standardisierten Beobachtung** festlegen. Die standardisierte Beobachtung erleichtert die Vergleichbarkeit und Quantifizierung der Informationen, engt das Beobachtungsfeld aber auch ein.

b) teilnehmende – nicht-teilnehmende

Bei der **teilnehmenden Beobachtung** übernimmt der Beobachter unter den Personen, die er beobachten soll, eine mithandelnde soziale Rolle. Die Möglichkeiten der teilnehmenden Beobachtung sind umso besser, je aktiver die zu beobachtende Gruppe ist.

Bei der **nicht-teilnehmenden Beobachtung** registriert der Beobachter „von außen her“. Er ist selbst nicht an den Handlungsabläufen beteiligt.

c) anekdotische – systematische

Bei der **anekdotischen Beobachtung** werden nur im Sinne der Fragestellung wichtige Ereignisse registriert, egal wann sie gerade auftreten. Diese Form der Beobachtung wird meist bevorzugt, wenn der Beobachter für längere Zeit am Leben der Beobachteten teilnimmt.

Bei der **systematischen Beobachtung** greift man häufig auf Zweitstichproben zurück. In festgelegter Reihenfolge werden die einzelnen Gruppenmitglieder beobachtet. Die Anzahl der beteiligten Personen richtet sich nach dem Beobachtungsziel. Es bestimmt auch, wie lange die Zeiträume sein müssen, um Verhaltensabläufe nicht zu sehr zu zerreißen.

Merkmals- (Zeichen-) und Kategoriensysteme

Der Beobachter muss vorher im Rahmen seiner Fragestellung festlegen, welche „Indikatoren“ er beobachten und wie er sie werten will. Er muss **Merkmale** oder **Kategorien** für seine Beobachtung bestimmen. Die Güte der Beobachtung hängt in hohem Maße von der Genauigkeit und Güte der Indikatoren ab.

a) Merkmalssystem

Man nimmt an, dass bestimmte beobachtbare Verhaltensweisen für einen ganzen Verhaltensbereich symptomatisch sind. Diese spezifischen Verhaltensaussagen werden im Voraus beschrieben. Ein Beobachter registriert dann, welche Verhaltensaussagen (z.T. auch wie oft) in einem Beobachtungszeitraum auftreten.

Beispiel: Oscar

Merkmalssysteme sind häufig „offen“, d.h. es können weitere Verhaltensaussagen angeführt werden, wenn nachgewiesen wird, dass die ausgewählten Merkmale besonders repräsentativ für die interessierende Fragestellung sind.

Oft ist anhand der Kategorien nicht sofort erkennbar, dass manche Verhaltensaussagen auch gleichzeitig auftreten können.

b) Kategoriensystem

SCHULZ 1973: „*Ein Kategoriensystem unterscheidet sich von einem Merkmalssystem dadurch, dass es Verhaltensweisen der protokollierten Art vollständig registrieren soll*“ (d.h., dass es außer den im System enthaltenen Kategorien keine mehr zu diesem Verhaltensbereich gibt). Es ist auch nicht möglich, dass zum gleichen Beobachtungszeitpunkt zwei Kategorien registriert werden können.

Die Anzahl der Kategorien sollte relativ klein sein: jede Kategorie wird somit ziemlich oft zur Registrierung einer relativ häufigen Verhaltensweise oder eines Verhaltensaspektes herangezogen. Auch für einen nicht in die Einzelheiten der Untersuchung Eingeweihten sollten die zu untersuchenden Verhaltensmerkmale ganz eindeutig sein.

Beispiel: Interaktionsprozess-Analysesystem nach Bales.

Kategoriensysteme zeichnen sich durch ein höheres Abstraktionsniveau aus, das gewöhnlich eine größere Anzahl von schlussfolgernden Schritten beim Beobachter voraussetzt. Merkmalssystem (Zeichensystem) erscheinen deshalb verlässlicher zu sein.

Das Problem der **Inferenz** kann auftreten. Inferenz bezeichnet die Wertungs- und Interpretationsprozesse, die sich an die Wahrnehmung des Phänomens anschließen

können. Urteile, die sich nicht auf vorher eindeutig festgelegte Beobachtungskategorien stützen, sind fast immer „hochinferente Urteile“. Das Problem der Inferenz versucht man m. H. des zusätzlichen Einsatzes von Schätzskaleten zu umgehen.

Schätzskaleten (Ratingskaleten)

Ratingskaleten werden im Vergleich zu Merkmals- und Kategoriensystemen (spät) nach der eigentlichen Beobachtung angewandt.

Neben der möglichen *zeitlichen Distanz* besteht der wesentliche Unterschied darin, dass nicht das Vorkommen eines Merkmals registriert wird, sondern eingestuft werden soll, *in welchem Grad* ein Merkmal oder eine Eigenschaft vorhanden ist.

Dem Problem der **Inferenz** begegnet man insofern, dass höher inferente Urteile erst nach mehreren niedriger inferenten Beobachtungen erreicht werden sollen.

Vielfach werden **Beurteilungshilfen** herangezogen. Als weitere Möglichkeit, die Beobachtung zu verbessern, wird die **Vorgabe operationalisierter Beobachtungshinweise** genannt. Man konzentriert sich dabei weniger auf einzelne Eigenschaften, sondern formuliert die Beobachtungsergebnisse auf einer Situationen-Verhaltensebene.

Diese Hilfen und Vorgaben sind unverzichtbar, wenn man die Objektivität, Zuverlässigkeit und Gültigkeit von Beobachtungen erhöhen will.

Allgemeine typische Fehler

- **Referenzfehler:** entstehen, wenn eine Beobachtung mit einem unangemessenen Muster verglichen wird („Nicht einordnen können oder wollen“)
- **Zusammenhangsfehler:** treten dann auf, wenn wir Merkmale in Verbindung bringen, die nicht notwendig in Zusammenhang stehen → Halo-Effekt („Verblendet sein“)

Gefahren für die Gültigkeit

Wahrnehmung ist nie ein Vorgang, bei dem wir neutral etwas „an sich“ wahrnehmen. Unsere Wahrnehmung ist immer durch physische, psychische und soziale Einflüsse gefärbt. Außerdem passen wir die Wahrnehmungsreize unseren Bedürfnissen und Erfahrungen an.

Die Hypothesentheorie der sozialen Wahrnehmung (BRUNER&POSTMAN 1951)

- **Selektion:** Wenn wir aus der Fülle der vorhandenen Reize diejenigen auswählen, die unserer Erwartung und unseren Bedürfnissen entsprechen
- **Organisation:** Wir strukturieren Wahrnehmungsreize so um, dass sie zu gewissen Persönlichkeitstheorien (Annahmen/Vorstellungen) oder Stereotypen passen
- **Akzentuierung:** Wir verleihen bestimmten Reizen besonderes Gewicht, während wir andere unterdrücken
- **Fixierung:** Wir geben einer Tendenz nach, die sich gegen Veränderung wehrt und einmal gewonnene Eindrücke auch dann auf neue Wahrnehmungsreize überträgt, wenn sie nicht dazu passen

Pädagogen werden meist auf nicht-standardisierte, teilnehmende und anekdotische Beobachtung zurückgreifen müssen. Das schränkt die Generalisierbarkeit der Beobachtungsergebnisse ein.

Befragungsmethoden

Allgemeine Merkmale und Definition

Durch Befragung können objektive (Fakten, Angaben zur Person) und subjektive (Meinungen, Einstellungen) Daten erhoben werden.

MANGOLD 1969 unterscheidet drei Arten von Informationen:

- Informationen über das *durchschnittliche (frühere, gegenwärtige oder zukünftig zu erwartende) Verhalten* bestimmter Bevölkerungsgruppen in bestimmten sozialen Situationen
- Informationen über die „subjektiven“ Bestimmungsgründe dieses Verhaltens (z.B. über *Meinungen, Motive, Verhaltensdispositionen und Erwartungen*)
- Informationen über „objektive“ soziale Gegebenheiten in der Situation der Befragten, die ihr *Verhalten ebenso wie ihr Bewusstsein* (und Unterbewusstsein) bestimmen

Befragungsmethoden werden meist folgendermaßen kategorisiert:

- mündliche oder schriftliche Erhebungstechnik
- Grad der Standardisierung
- Individual- oder Gruppenerhebung
- Einmalige oder mehrfache Anwendung

SCHEUCH 1973 definiert: „Unter Interview als Forschungsinstrument sei hier verstanden ein planmäßiges Vorgehen mit wissenschaftlicher Zielsetzung, bei dem die Versuchsperson durch eine Reihe gezielter Fragen oder Stimuli zu verbalen Informationen veranlasst werden sollen“

Im Interview sind offene und geschlossene Fragen möglich:

offene Fragen	geschlossene Fragen
<ul style="list-style-type: none">○ können den Gesprächskontakt und das Interesse am Interview fördern, weil sie einer normalen Gesprächssituation nahe kommen○ besonders geeignet, um im Planungsstadium das Problemfeld zu erforschen und die relevanten Antwortkategorien zu erfassen	<ul style="list-style-type: none">○ erbringen eine größere Einheitlichkeit der Antworten und erleichtern dadurch die Vergleichbarkeit○ geeignet zur Überprüfung von Hypothesen

Interview-Typen

Nicht standardisiertes Interview

Das nicht standardisierte Interview hat meist den Zweck einer *ersten Erkundung*.

Der Befragung wird lediglich ein Rahmenthema gegeben, über das sich der Interviewer mit dem Befragten unterhält. Der Interviewer steuert das Gespräch höchstens durch Zwischenfragen. Der Befragte hat die optimale Möglichkeit, das Interview so zu gestalten, dass er seine Einstellungen und Interessen so ausdrücken kann, wie er es will. Diese freie Form stellt *sehr hohe Anforderungen* an den Interviewer.

Standardisierte Interviews

Beim standardisierten Interview ist der Fragesteller an ein ausgearbeitetes Schema gebunden, in dem der Ort des Interviews, die Formulierung der Fragen als auch ihre Reihenfolge und die möglichen Erläuterungen zu den Fragen festgelegt sind.

In der Regel ist die Reliabilität bei standardisierten Interviews größer. Fehler bei der Formulierung der Frage werden auf ein Minimum reduziert.

Schriftliche Befragung oder Fragebogen

Die schriftliche Befragung ist ein deutlich weniger aufwändiges Verfahren, um Auskünfte über objektive Daten und subjektive Einstellungen zu erlangen.

Das Vorgehen bei der Konstruktion eines Fragebogens beschreibt SCHEUCH 1969:

- 1) Umsetzung eines Themas in einzelne Erhebungspunkte
- 2) Entwurf der Fragen (offene Fragen sind in dieser Phase üblich)
- 3) Vorstellung über die Art der Auswertung dieser Fragen
- 4) Durchführen von Vortests und Probebefragungen mit besonders geschulten Interviewern
- 5) Quantitative Überprüfung, meist unter Verwendung „gegabelter Fragebogen“
- 6) Durchführen der eigentlichen Befragung

Ein neueres Verfahren beschreibt MUMMENDEY:

- 1) Bestimmung der Form des Fragebogens (Art d. Itemformulierung, d. Antworttypus)
- 2) Itemsammlung (Theoriegeleitet, Erfahrungsbezogen, Literatur, etc.)
- 3) Itemrevision (Verständlichkeit muss gegeben sein)
- 4) Fragebogeninstruktion (Abfassung der Antwortanweisungen)
- 5) Itemanalyse (Schwierigkeit, Trennschärfe)
- 6) Reliabilität (interne Konsistenz, Stabilität)
- 7) Validität (Inhalts-Val., Val. als Korrelation mit Kriterium, Konstrukt-Val.)

Die diagnostische Zielsetzung einzelner Pädagogen ist angesichts dieses Konstruktionsaufwandes sehr viel eingegrenzter und hat einen anderen Anspruch. Dabei werden vorwiegend offene Fragen anonym beantwortet.

Tests

Definition und Klassifikation

„Tests sind Verfahren der Pädagogischen Diagnostik, mit deren Hilfe eine Verhaltensstichprobe, die Voraussetzungen für oder Ergebnisse von Lernprozessen repräsentieren soll, möglichst vergleichbar, objektiv, zuverlässig und gültig gemessen und durch Lehrer oder Erzieher ausgewertet, interpretiert und für ihr pädagogisches Handeln nutzbar gemacht werden kann“

BRICKENKAMP 1975 unterscheidet Tests nach dem Untersuchungsgebiet:

- a) Leistungstests („Gib Dein Bestes!“)
- b) Psychometrische Persönlichkeitstests („Wie bist Du durchschnittlich?“)

Tests legen von der Testvorbereitung über Aufgabendarbietung, -auswertung bis zur Interpretation alle Durchführungsschritte fest, um möglichst vergleichbare und objektive Ergebnisse zu erhalten.

Testpräsentation

Im pädagogischen Bereich herrschen Gruppentests vor. Individualtests werden in der Regel nur durchgeführt, wenn das geforderte Verhalten eine Gruppenprüfung unmöglich macht.

Im Prinzip handelt es sich dabei um sog. Staffel- bzw. Leistungstests, die auf der jeweiligen Leistungsstufe des zu Testenden ansetzen und dann schrittweise gesteigert werden.

Sequentiell antwortabhängiges Testen

- homogener, im Prinzip nicht begrenzter Aufgabepool
- Reihenfolge der Aufgaben von der Lösung der zuletzt vorgelegten Aufgabe abhängig
- Durchführung so lange, bis der durch Experten festgelegte Beherrschungsgrad erreicht ist

Veranschaulichung durch Schema von SIXTL 1978

Rangmäßige Einstufungen sind durch den Test nicht möglich. Es kann nur ausgesagt werden „Ziel erreicht“ oder „Ziel nicht erreicht“.

Adaptives antwortabhängiges Testen

Adaptives Testen heißt „maßgeschneidertes“ Testen. Man orientiert sich am Leistungsniveau des Probanden. Auf diese Art lässt sich das individuelle Leistungsniveau differenzierter ermitteln. Diese Form des Testens wird häufig Computergestützt durchgeführt.

- zusätzlich zur Aufgabenzahl wird die Aufgabenqualität berücksichtigt
- die Testperson bearbeitet nur die für sie relevanten Aufgaben
- Welche Stichprobe die Testperson bearbeitet hängt also ganz von ihren Aufgabenlösungen ab

Formen der Testaufgabe

RÜTTER unterscheidet drei Klassen von Testaufgaben:

- a) offene Testaufgaben

- b) halboffenen Testaufgaben
- c) geschlossene Testaufgaben

Für praktische Zwecke genügt auch eine Unterscheidung in die **Freie Antwort-Form** und **Auswahl-Form**:

Freie Antwort-Formen

- Lückentexte
- Ergänzungsaufgaben
- Kurzantwortform
- Kurzaufsatzform

Auswahl-Formen

- Zuordnungsform
- Alternativformen
- Antwort-Auswahl-Form

Klassische oder probabilistische Testtheorie?

Klassische Testtheorie

Als formale, normative Theorie beschäftigt sie sich nicht mit dem Zustandekommen eines Messwertes oder der Beziehung von Merkmal und Messwert. Es geht um die Frage nach dem Ausmaß von verfälschenden Anteilen bei Messungen. Man spricht deshalb auch von der „Messfehlertheorie“.

Die wichtigsten Voraussetzungen der KTT wurden von HELLER 1974 in drei Axiomen zusammengefasst:

1. Existenzaxiom
2. Fehleraxiom
3. Verknüpfungsaxiom

„Neue (probabilistische) Testtheorie“

Die PTT fasst den beobachteten Wert lediglich als Indikator eines latenten (nicht beobachtbaren) Merkmals auf, das mit dem Instrument gemessen werden soll. Man spricht deshalb auch von der „Latent-Trait-Theorie“

In der Praxis hat sich die PTT noch nicht durchgesetzt.

Literatur*

INGENKAMP, K (1988): Lehrbuch der Pädagogischen Diagnostik. Weinheim

LIENERT, G.A. (1969): Testaufbau und Testanalyse. Weinheim

BRICKENKAMP, R (Hrsg.) (1997): Handbuch psychologischer und pädagogischer Tests. Göttingen²

Nicht vollständig