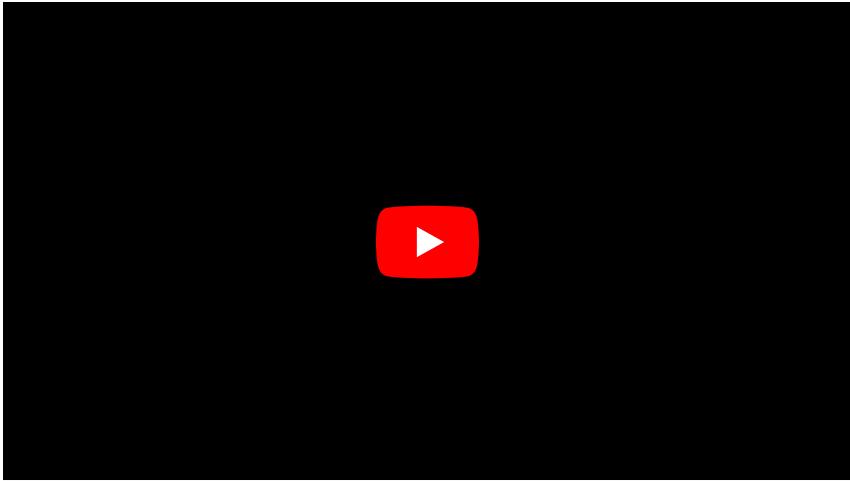


Einführung in die Forschungsmethoden der Psychologie und Psychotherapie

Einheit 3: Manifeste Merkmale: Verhaltensbeobachtung

25.04.2024 | Dr. Caroline Zygar-Hoffmann

Verhaltensbeobachtung



| Was beobachten Sie?

Verhaltensbeobachtung

- Alltagsbeobachtung vs. wissenschaftliche Beobachtung
- Beobachtungsplan
- Beobachtungsfehler und -verzerrungen
- Beobachterübereinstimmung

Take-Aways und Schlüssel-/Fachbegriffe

- Take-Aways
- Schlüssel-/Fachbegriffe

Alltagsbeobachtung vs. wissenschaftliche Beobachtung

Der Alltag aber auch jede Datenerhebung beinhaltet den Vorgang der visuellen Wahrnehmung, also Beobachtung (aber visuelle Wahrnehmung \neq wissenschaftliche Beobachtung)

Herausforderungen für Beobachtungen:

- **Perspektivität**
 - Ziel, Perspektive und Standpunkt beeinflussen die Beobachtung
- **Selektivität**
 - Informationen müssen aus komplexem, sich ständig veränderndem Strom von ständig eintreffenden visuellen Daten ausgewählt werden
- **Konstruktivität**
 - Selektierte Informationen müssen zu bedeutungsvollen, stabilen Einheiten zusammengefügt werden
 - Konstruktion abhängig von Einstellungen, Erwartungen und Vorerfahrungen der Beobachter:in

Lösung: Wissenschaftliche Beobachtung nach a priori festgelegtem Beobachtungsplan, um intersubjektiv nachvollziehbare Erkenntnisse zu gewinnen

Alltagsbeobachtung vs. wissenschaftliche Beobachtung

- Verhalten ist erstmal direkt beobachtbar, und kann dadurch direkt als manifeste Variable ausgewertet werden.
- Verhaltensbeobachtung kann aber auch zur Messung von latenten Variablen dienen (meist durch eine darauffolgende Verhaltensbeurteilung).

Annahmen der Verhaltensbeobachtung zur Messung latenter Variablen:

1. **Typisches Verhalten** einer Person wird durch **Eigenschaften der Person** bewirkt.
2. Eine höhere Ausprägung einer bestimmten Eigenschaft macht das Auftreten des von ihr bewirkten Verhaltens wahrscheinlicher.
3. Es gibt interindividuelle Unterschiede im Verhalten.
4. Beobachtungsraum = „Bildausschnitt“ (Verhaltensstichprobe)
5. Interindividuelle Unterschiede sind über die Zeit hinweg stabil.
6. Interindividuelle Unterschiede sind über verschiedene Situationen hinweg konsistent.

→ Gewohnheiten sind auf die Wirkung von einer oder mehreren Eigenschaften zurückzuführen

Alltagsbeobachtung vs. wissenschaftliche Beobachtung

Definition wissenschaftliche Beobachtung/Beurteilung: Zielgerichtete, systematische und regelgeleitete Erfassung, Dokumentation und Interpretation des Auftretens bzw. der Ausprägung von ausgewählten, psychologisch relevanten Merkmalen, Verhalten oder Ereignissen

"Verhaltensbeobachtung liefert im Idealfall Fakten, die frei von Bewertungen sind."

(Schmidt-Atzert & Amelang, 2012, S.309)

- Diese Zielsetzung impliziert eine wichtige Unterscheidung zwischen Beobachtung (Wahrnehmung und Registrierung) vs. Beurteilung
- Wahrnehmung, Registrierung (bzw. Protokollierung) und Beurteilung sollten im Prozess der Verhaltensbeobachtung getrennt vorgenommen werden
- Eine gute Wahrnehmung und Registrierung setzt voraus, dass genau definiert wird, was und wie etwas beobachtet (und später beurteilt) werden soll

Beobachtungsplan

Wissenschaftliche Beobachtung nach a priori festgelegtem Beobachtungsplan:

1. Wie wird beobachtet? (Grad der Strukturierung)
2. Was wird beobachtet? (Festlegung von Beobachtungseinheiten = Indikatoren)
3. Wer beobachtet? (teilnehmend vs. nicht-teilnehmend)
4. Wo wird beobachtet? (Feld vs. Labor)
5. Wann wird beobachtet und wie wird Beobachtung wahrgenommen? (Zeitpunkte, technische Vermittlung)
6. Wie wird Beobachtung registriert/protokolliert?
7. Wie wird Beobachtung beurteilt?

Beobachtungsplan

Formen und Rahmenbedingungen für Beobachtungen

Wie wird beobachtet? - Strukturierte vs. unstrukturierte Beobachtung:

Strukturierte Beobachtung wird anhand eines vorher festgelegten Beobachtungsplans durchgeführt

- höhere Objektivität und Reliabilität
- Quantifizierung des beobachteten Verhaltens möglich

Unstrukturierte Beobachtung findet ohne genauen Beobachtungsplan statt

- exploratives Vorgehen wenn zu wenig Informationen zur Erstellung eines Beobachtungsplans vorliegen
- meist im Rahmen von qualitativer Forschung
- Beispiel: Stanford Prison Experiment (Haney et al., 1973; Zimbardo & Haney, 2008; welches aber aus heutiger Perspektive wegen methodischer Mängel stark kritisiert wird, vgl. Le Texier, 2019)

Beobachtungsplan

Beobachtungseinheiten

Was wird beobachtet? - Definition von Beobachtungseinheiten = Indikatoren:

- Abgrenzung des Beobachtungsobjekts: z.B. einzelne Personen vs. Familie vs. Paare...
- Abgrenzung der Variablen / Verhaltensmerkmale: umfassendes Bild vs. einzelnes Merkmal
- Abgrenzung der Dauer der Beobachtung: kurz/einmalig vs. lang/wiederholt
- Definition von Anfang, Ende und Inhalt einer Beobachtungseinheit:
 - Wann beginnt / endet ein Verhalten?
 - Welches Verhalten gibt Hinweis auf latentes Merkmal?

Beobachtungsplan

Beobachtungseinheiten

Was wird beobachtet? - Kriterien für gute Beobachtungseinheiten = Indikatoren:

- Alle Beobachtende **verstehen** unter den Beobachtungseinheiten **dasselbe** → verhaltensnah, konkret bzw. genau definiert (Objektivität)
- Beobachtungseinheiten können klar voneinander **abgegrenzt** werden („unterbricht andere“ vs. „redet dazwischen“) (Objektivität)
- Das beschriebene Verhalten **kommt** in der Beobachtungssituation **vor** (Bachtung der Situationsabhängigkeit für Validität der Messung)
- Jeder Beobachtende kann Beobachtungseinheit **zuverlässig identifizieren** (Reliabilität).
- Liste der Beobachtungseinheiten kann von Beobachtenden während der Beobachtung **überblickt** werden (Reliabilität). **Grenzwerte** für Beobachtende: max. 3 Eigenschaften, max. 10 Einzelverhaltensweisen, max. 20 Minuten pro Beobachtung (abhängig von Erfahrung der Beobachtenden)
- Zu beobachtendes Merkmal wird von Beobachtungseinheiten in **allen Facetten** erfasst (Inhaltsvalidität).

Tipp: Beobachtungsplan in einem „Probedurchgang“ testen

Beobachtungsplan

Beobachtungseinheiten

Kategoriensystem: Gliederung des beobachtbaren Verhaltens in Kategorien als Indikatoren

"Mithilfe von Kategoriensystemen will man ein Verhalten vollständig anstatt über ausgewählte Verhaltensweisen erfassen. Das Verhalten wird dazu in mehrere klar definierte und voneinander abgrenzbare Kategorien unterteilt. Die Kriterien Vollständigkeit, Eindeutigkeit und Überschneidungsfreiheit sind jedoch schwer zu erfüllen; Kategoriensysteme spielen daher in der diagnostischen Praxis praktisch keine Rolle."

(Schmidt-Atzert & Amelang, 2012, S.309)

Beispiel: Aggressives Verhalten von Kindergartenkindern

1. schlägt, tritt
2. schlägt zurück
3. wirft mit Gegenstand
4. wirft mit Gegenstand zurück
5. droht, beschimpft
6. droht, schimpft zurück
7. zerstört beim Partner
8. zerstört Eigenes
9. zerstört beim Partner als Reaktion
10. zerstört Eigenes als Reaktion

Beobachtungsplan

Formen und Rahmenbedingungen für Beobachtungen

Wer beobachtet? - Teilnehmende vs. nicht-teilnehmende Beobachtung:

Aktiv-teilnehmende Beobachtung vs. passiv-teilnehmende Beobachtung

- keine Geräte für Beobachtung notwendig
- bei passiv-teilnehmender Beobachtung: stummer Beobachter
- bei aktiv-teilnehmender Beobachtung: Beobachter Teil der Situation → "natürliche" Situation möglich

Nicht-teilnehmende Beobachtung:

- Beobachtende sind in der Situation nicht anwesend bzw. werden von Beobachtetem nicht gesehen, z.B. per Kamera, einseitig durchsichtige Scheibe
- Situation wird nicht so stark verändert

ggf. Anschlussfragen: verdeckte vs. offene Beobachtung (siehe Einheit zu qualitativen Methoden); Selbstbeobachtung vs. Fremdbeobachtung

Beobachtungsplan

Formen und Rahmenbedingungen für Beobachtungen

Wo wird beobachtet? - Feld- vs. Laborbeobachtung:

Bei Beobachtung im Feld wird in der „natürlichen“ Situation beobachtet

- Situation ist „echt“
- höhere **ökologische Validität** (Übertragbarkeit auf echtes Verhalten)

Bei Beobachtung im Labor wird in einer „künstlich erzeugten“ Situation beobachtet. Rahmen- bedingungen und Aufgaben sind genau geplant

- Kontrolle von Störfaktoren
- höhere Vergleichbarkeit

Beispiele für Laborbeobachtungen:

- Fremde-Situations-Test zum Bindungsverhalten von Kindern, mit klarem Ablauf wann Bezugsperson den Raum verlässt, fremde Person hinzukommt etc. (Ainsworth & Wittig, 1969)
- Bandura's Bobo Doll ([Link](#)), eine der wichtigsten Studien zum Beobachtungslernen aka Modell-Lernen
- Das Marshmallow Experiment (Mischel, 1989) ([Link](#)): Experiment zum Belohnungsaufschub und zur Impulskontrolle

Beobachtungsplan

Formen und Rahmenbedingungen für Beobachtungen

Wann wird beobachtet? - Direkte (gleichzeitige) vs. indirekte (zeitversetzte) Beobachtung und damit z.T. verbunden: **Wie wird Beobachtung wahrgenommen? - Vermittelte vs. unvermittelte Beobachtung**

Direkte Beobachtung: Verhalten wird direkt zum Zeitpunkt des Auftretens beobachtet

- entweder vermittelt durch eine Kamera oder unvermittelt
- zielgerichtete Wahrnehmung auf bestimmte Aspekte der Situation durch Repositionierung des Beobachters oder Repositionierung einer Kamera möglich
- Vorteil einer unvermittelten, direkten Beobachtung: keine Geräte notwendig, in allen Situationen möglich

Indirekte Beobachtung: zeitversetzte „Beobachtung“

- setzt eine per Kamera vermittelte Beobachtung voraus, d.h. nachträgliche Verhaltensbeobachtung von auf Video aufgezeichnetem Verhalten
- längere Verhaltensausschnitte möglich, Fehlerminimierung durch wiederholtes Abspielen
- hierunter kann auch eine Dokumentenanalyse fallen, zur Beobachtung von Verhaltensspuren in jeglicher Art von untersuchbarem Material (von Akten über Briefen bis zu Fotos) → siehe Einheit zu Qualitativen Methoden

Beobachtungsplan

Formen und Rahmenbedingungen für Beobachtungen

Wie wird Beobachtung registriert? - Time vs. Event Sampling

Event Sampling (Ereignistichprobe):

Registrierung von Häufigkeit und/oder Dauer des Verhaltens- oder Situationsmerkmals **über den gesamten Beobachtungszeitraum hinweg**

- es werden eher auch seltene Ereignisse erfasst, weil es keine Unterbrechungen gibt
- erlaubt bessere Überprüfung von Hypothesen zum Zusammenhang aufeinanderfolgender Ereignisse (Kontingenzen) als Time Sampling

Time Sampling (Zeitstichprobe):

- Unterteilung des Beobachtungszeitraums in gleich lange Zeitintervalle
- Registrierung ob/wie lange Verhaltensmerkmal im Zeitintervall gezeigt wird
- Möglichkeit, Pausen einzubauen (z.B. nur in jedem 2. Zeitintervall beobachten)
- Beobachterrotation möglich (z.B. Beobachtender 1 im 1. Zeitintervall, Beobachtender 2 im 2. Zeitintervall)
- oft akustischer oder vibrierender Signalgeber notwendig

Verhaltensbeobachtung

Beobachtungsplan

Formen und Rahmenbedingungen für Beobachtungen

Wie wird Beobachtung registriert? - Beispiel für Time Sampling

Indikator / Merkmalsbereich	Zeitintervall (1 bis n)						Σ
	1	2	3	4	...	n	
Greift Vorschlag anderer auf							
Revidiert eigenen Vorschlag							
Zeigt Zustimmung							
Σ Kooperation							
Verteidigt eigenen Vorschlag							
Kritisiert andere							
Unterbricht andere							
Σ Durchsetzung							

Beobachtungsplan

Formen und Rahmenbedingungen für Beobachtungen

Wie wird Beobachtung beurteilt? - Konstruktion von Ratingskalen

- unipolare vs. bipolare Skala
- Anzahl der Abstufungen: abhängig von Differenzierungsfähigkeit der Beobachter:in und Variation des Merkmals
- Markierung und verbale Verankerung der Skalen: um Abstufungen der Skalen deutlich zu machen, sollten diese numerisch/grafisch markiert und verbal verankert werden
- Viele Überschneidungen zur Konstruktion von Ratingskalen bei Befragungen (siehe nächste Einheit)

unipolar:

|-----| feindselig

numerisch verankert:

0 - 1 - 2 - 3

bipolar:

freundlich |-----| feindselig

verbal verankert:

sehr niedrig - eher niedrig - eher hoch - sehr hoch

Beobachtungsplan

Formen und Rahmenbedingungen für Beobachtungen

Wie wird Beobachtung beurteilt? - Konstruktion von Ratingskalen

- Ratingskalen für eine Verhaltensbeobachtung sollten zusätzlich grundsätzlich um Verhaltensanker ergänzt werden → **verhaltensverankerte Ratingsskala**

Aggressivität				
0	1	2	3	4
gar nicht	wenig	etwas	deutlich	stark
0 = zeigt kein verbales oder körperlich aggressives Verhalten 1 = schimpft gelegentlich zurück, wird aber nicht initiativ 2 = schimpft, wenn etwas nicht klappt 3 = schreit andere an oder wirft Sachen um 4 = brüllt bei geringsten Anlässen oder wirft mit Gegenständen oder rempelt andere an				

- Alternativ braucht es klare Regeln zur Zuordnung zwischen den beobachteten Indikatoren und der Zuordnung auf der Ratingskala

Beobachtungsfehler und -verzerrungen (Beispiele)

Beobachtungsfehler	Beschreibung
Beobachterdrift	Im Verlauf zunehmend weniger/mehr Auftreten registrieren, obwohl sich am Verhalten nichts verändert hat (oft wegen Ermüdung, mangelnder Übung)
Erinnerungseffekte	Verzerrungen, wenn nicht unmittelbar protokolliert wird
Haloeffekt	Eine Eigenschaft einer Person wird durch eine andere Eigenschaft „überstrahlt“
Logischer Fehler	Von einem Hinweisreiz wird auf eine Eigenschaft geschlossen (obwohl kein Zusammenhang besteht)
Primacy Effekt	Überbetonung der ersten Information
Recency Effekt	Überbetonung der letzten Information
Mildeeffekt	Verzerrung zum Positiven
Strengeeffekt	Verzerrung zum Negativen
Zentrale Tendenz	Bevorzugung der Skalenmitte (häufig wegen Unsicherheit beim Beurteilen)
Tendenz zu Extremurteilen	Bevorzugung der Skalenpole
Kontrasteffekt	Einschätzung einer Person hängt von jener der zuvor eingeschätzten Person ab

Verringern von Beobachtungsfehlern und -verzerrungen

Maßnahmen	Beschreibung
Grenzwerte	Limit an zu beobachtenden Verhaltensmerkmalen und Beobachtungsdauer für BeobachterInnen
Verhaltensahe Beobachtung	Beobachtbare Einheiten, Vermeidung von Interpretation
konkreter/eindeutiger Beobachtungsplan	Missverständnisse vermeiden, klare Kategorien
Beobachterschulung	Vertrautmachen mit Protokollbogen, Definieren der Verhaltensmerkmale, Probedurchgänge
Beobachterrotation	verschiedene Beobachter:innen einsetzen; Versuchsperson in verschiedenen Situationen durch unterschiedliche Beobachter:innen beobachten

- Auch bei Berücksichtigung von Maßnahmen sind systematische und unsystematische Beobachtungsfehler nicht auszuschließen
- Daher bei jeder Verhaltensbeobachtung Beobachter- bzw. Beurteilerübereinstimmung empirisch prüfen!
- Trotzdem: auch Beobachterübereinstimmung garantiert noch nicht, dass keine systematischen Beobachtungsfehler aufgetreten sind
- Mittelung über die Beobachtungen mehrerer Beobachter:innen verbessert die Reliabilität der Messung (da zumindest unsystematische Messfehler rausgemittelt werden)

Maße für die Beobachter- und Beurteilerübereinstimmung

- Um Beobachterübereinstimmung und Reliabilität einer Verhaltensbeobachtung zu bestimmen gibt es eine Vielzahl von Maßen
- Wahl des angemessenen Übereinstimmungs- bzw. Zuverlässigkeitsmaßes abhängig vom Skalenniveau der Daten
- Die wichtigsten Maße (lernen Sie in der Vorlesung zu psychologischer Diagnostik im 3.Semester genauer kennen):
 - prozentuale Übereinstimmung
 - Cohen's Kappa oder Fleiss' Kappa (mindestens nominale Daten)
 - Kendalls Konkordanzkoeffizient (mindestens ordinale Daten)
 - Intra-Class-Korrelation (ICC) (metrische Daten)

Maße für die Beobachterübereinstimmung

Beispiel: nominales Urteil (ja/nein)

Vehalten	Rater1	Rater2
1. schlägt, tritt	Ja	Nein
2. schlägt zurück	Ja	Ja
3. wirft mit Gegenstand	Nein	Nein
4. wirft mit Gegenstand zurück	Ja	Ja
5. droht, beschimpft	Ja	Ja
6. droht, schimpft zurück	Ja	Ja
7. zerstört beim Partner	Nein	Ja
8. zerstört Eigenes	Ja	Ja
9. zerstört beim Partner als Reaktion	Nein	Nein
10. zerstört Eigenes als Reaktion	Nein	Nein

Maße für die Beobachterübereinstimmung

Beispiel: metrisches Urteil (Skala 1-6)

Vehalten	Rater1	Rater2
1. schlägt, tritt	3	4
2. schlägt zurück	3	2
3. wirft mit Gegenstand	5	5
4. wirft mit Gegenstand zurück	1	2
5. droht, beschimpft	4	5
6. droht, schimpft zurück	1	1
7. zerstört beim Partner	6	6
8. zerstört Eigenes	5	3
9. zerstört beim Partner als Reaktion	2	2
10. zerstört Eigenes als Reaktion	1	1

Verhaltensbeobachtung

Übungsaufgabe:

Überlegen Sie sich geeignete Beobachtungseinheiten bei einer Event Sampling Beobachtung (wie oft wird Verhalten gezeigt) für **Werbungsverhalten von Säuglingen um die Aufmerksamkeit ihrer Eltern** (in der Beobachtungssituation sitzt das Baby in einer Wippe neben den Eltern)

[Link zur Beobachtungssituation](#)

Fragen zum realisierten Beobachtungsplan:

- teilnehmend vs. nicht-teilnehmend?
- Feld vs. Labor?
- Direkt vs. indirekt?
- Vermittelt vs. unvermittelt?

Überlegen Sie sich, inwiefern Sie die Hauptgütekriterien für Ihre Verhaltensbeobachtung erfüllt sehen. Wie würden Sie das jeweils überprüfen?

- Wissenschaftliche Beobachtung erfolgt nach a priori festgelegtem **Beobachtungsplan**
- **Indikatoren** sollten verhaltensnah, genau definiert, abgegrenzt, situationsbezogen, zuverlässig identifizierbar und überblickbar sein
- Für **Inhaltsvalidität** muss die Gesamtheit der Indikatoren betrachtet werden und ob sie ein Merkmal hinreichend abdecken
- Eine **Registrierung per Event Sampling oder Time Sampling** hat verschiedene Vor- und Nachteile
- **Verhaltensverankerte Ratingskalen** sind unerlässig für eine gute Beurteilung
- **Beobachtungsfehler und -verzerrungen** sollten möglichst bewusst vermieden werden
- Zur Bestimmung der Objektivität und Reliabilität von Beobachtungen können **Maße der Beobachterübereinstimmung** berechnet werden

Schlüssel-/Fachbegriffe der heutigen Vorlesung

Perspektivität	aktiv-teilnehmende Beobachtung	unvermittelte Beobachtung	Haloeffekt
Selektivität	passiv-teilnehmende Beobachtung	Time Sampling	Logischer Fehler
Konstruktivität	nicht-teilnehmende Beobachtung	Event Sampling	Primacy Effekt
strukturierte Beobachtung	direkte Beobachtung	unipolare Skala	Recency Effekt
unstrukturierte Beobachtung	indirekte Beobachtung	bipolare Skala	Kontrasteffekt
Indikator	vermittelte Beobachtung	verhaltensverankerte Ratingskala	Cohen's / Fleiss' Kappa
ökologische Validität		Beobachterdrift	Kendalls Konkordanzkoeffizient
			Intra-Class-Korrelation

[zurück zur heutigen Übersicht der Vorlesung →](#)

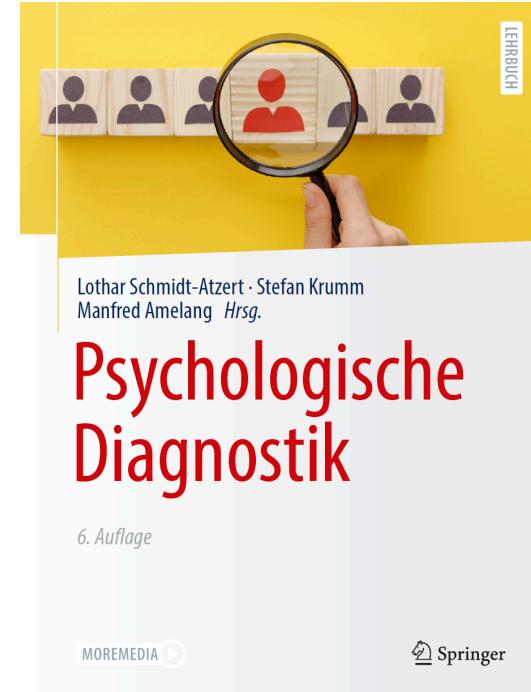
[zum Quiz zur Wissensprüfung →](#)

Literatur für die heutige Sitzung



Kapitel 10.1. in Döring, N. & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Pearson.

Materialien: Vielen Dank an Prof. Dr. Stephan Goerigk und den Lehrstuhl für Psychologische Methodenlehre und Diagnostik an der LMU für Bereitstellung der Grundlage für die Materialien



Kapitel 3.6. in Schmidt-Atzert, L., Krumm, S. & Amelang, M. (2021). *Psychologische Diagnostik*. Springer.