

Jürgen Bortz
Nicola Döring

Forschungsmethoden und Evaluation

für Human- und Sozialwissenschaftler

4., überarbeitete Auflage

Mit 156 Abbildungen und 87 Tabellen

Inhaltsverzeichnis

Zu diesem Buch	XVII	2	Von einer interessanten Fragestellung zur empirischen Untersuchung	35
1 Empirische Forschung im Überblick		1		
1.1 Begriffe und Regeln der empirischen Forschung		2	2.1 Themensuche	36
1.1.1 Variablen und Daten	2		2.1.1 Anlegen einer Ideensammlung	37
1.1.2 Alltagsvermutungen und wissenschaftliche Hypothesen	4		2.1.2 Replikation von Untersuchungen	37
Der Informationsgehalt von Wenn-dann-Sätzen			2.1.3 Mitarbeit an Forschungsprojekten	38
Wenn- und Dann-Teil als Ausprägungen von Variablen			2.1.4 Weitere Anregungen	38
Statistische Hypothesen			2.2 Bewertung von Untersuchungsideen	40
Prüfkriterien			2.2.1 Wissenschaftliche Kriterien	40
1.1.3 Kausale Hypothesen	11		Präzision der Problemformulierung	
Mono- und multikausale Erklärungen			Empirische Untersuchbarkeit	
Wenn-dann-Heuristik			Wissenschaftliche Tragweite	
Messfehler und Störvariablen			2.2.2 Ethische Kriterien	41
1.1.4 Theorien, Gesetze, Paradigmen	15		Güterabwägung: Wissenschaftlicher Fortschritt oder Menschenwürde	
1.2 Grenzen der empirischen Forschung		16	Persönliche Verantwortung	
1.2.1 Deduktiv-nomologische Erklärungen	16		2.2.3 Informationspflicht	44
1.2.2 Verifikation und Falsifikation	18		Freiwillige Untersuchungsteilnahme	
Korrespondenz- und Basissatzprobleme			Vermeidung psychischer oder körperlicher Beeinträchtigungen	
1.2.3 Exhaustion	21		Anonymität der Ergebnisse	
1.3 Praktisches Vorgehen		22	2.3 Untersuchungsplanung	46
1.3.1 Statistische Hypothesenprüfung	23		2.3.1 Zum Anspruch der geplanten Untersuchung	46
Untersuchungsplanung			2.3.2 Literaturstudium	47
Statistisches Hypothesenpaar			Orientierung	
Auswahl eines Signifikanztests			Vertiefung	
Das Stichprobenergebnis			Dokumentation	
Berechnung der Irrtumswahrscheinlichkeit mittels Signifikanztest			2.3.3 Wahl der Untersuchungsart	49
Signifikante und nicht signifikante Ergebnisse			Erstes Kriterium: Stand der Forschung	
Signifikanzniveau			Zweites Kriterium: Gültigkeit der Untersuchungsbefunde	
1.3.2 Erkenntnisgewinn durch statistische Hypothesentests?	27		2.3.4 Thema der Untersuchung	59
Das »Good-enough-Prinzip« – eine Modifikation des Signifikanztests			2.3.5 Begriffsdefinitionen und Operationalisierung	60
1.4 Aufgaben der empirischen Forschung		29	Real- und Nominaldefinitionen	
1.4.1 Hypothesenprüfung und Hypothesenerkundung	30		Analytische Definitionen	
1.4.2 Empirische Forschung und Alltagserfahrung	31		Operationale Definitionen	
Systematische Dokumentation			2.3.6 Messtheoretische Probleme	65
Präzise Terminologie			Was ist Messen?	
Statistische Analysen			Skalenarten	
Interne und externe Validität			Praktische Konsequenzen	
Umgang mit Theorien			2.3.7 Auswahl der Untersuchungsobjekte	70
Übungsaufgaben			Art und Größe der Stichprobe	
			Anwerbung von Untersuchungsteilnehmern	
			Determinanten der freiwilligen Untersuchungsteilnahme	
			Studierende als Versuchspersonen	
			Empfehlungen	

2.3.8	Durchführung, Auswertung und Planungsbericht	75	3.2.3	Operationalisierung von Maßnahme- wirkungen	116
	Planung der Untersuchungsdurchführung			Varianten für unabhängige Variablen	
	Aufbereitung der Daten			Erfassung der abhängigen Variablen	
	Planung der statistischen Hypothesenprüfung			Überlegungen zur Nutzenbestimmung	
	Interpretation möglicher Ergebnisse			Abstimmung von Maßnahme und Wirkung	
	Exposé und Gesamtplanung		3.2.4	Stichprobenauswahl	127
2.4	Theoretischer Teil der Arbeit	81		Interventionsstichprobe	
2.5	Durchführung der Untersuchung	81		Evaluationsstichprobe	
2.5.1	Versuchsleiterartefakte	82	3.2.5	Abstimmung von Intervention und Evaluation	130
2.5.2	Praktische Konsequenzen	83	3.2.6	Exposé und Arbeitsplan	131
2.5.3	Empfehlungen	83	3.3	Durchführung, Auswertung und Berichterstellung	132
2.6	Auswertung der Daten	85	3.3.1	Projektmanagement	132
2.7	Anfertigung des Untersuchungsberichtes	86	3.3.2	Ergebnisbericht	132
2.7.1	Gliederung und Inhaltsverzeichnis	86	3.3.3	Evaluationsnutzung und Metaevaluation	133
2.7.2	Die Hauptbereiche des Textes	87	3.4	Hinweise	134
	Abstract			Übungsaufgaben	
	Einleitung		4	Quantitative Methoden der Datenerhebung	137
	Forschungsstand und Theorie		4.1	Zählen	139
	Methode		4.1.1	Qualitative Merkmale	140
	Ergebnisse		4.1.2	Quantitative Merkmale	143
	Diskussion		4.1.3	Indexbildung	143
	Literatur			Auswahl und Art der Indikatoren	
2.7.3	Gestaltung des Manuskripts	90		Zusammenfassung der Indikatoren	
2.7.4	Literaturhinweise und Literaturverzeichnis	90		Gewichtung der Indikatoren	
2.7.5	Veröffentlichungen	93		Index als standardisierter Wert	
	Übungsaufgaben		4.1.4	Quantitative Inhaltsanalyse	149
3	Besonderheiten der Evaluationsforschung	95		Geschichte der Inhaltsanalyse	
3.1	Evaluationsforschung im Überblick	96		Anwendungsfelder	
3.1.1	Evaluationsforschung und Grundlagen- forschung	98		Das Kategoriensystem	
	Gebundene und offene Forschungsziele			Die Textstichprobe	
	Entscheidungszwänge und wissenschaftliche Vorsicht			Kodierung und Kodierereinheit	
	Technologische und wissenschaftliche Theorien			Statistische Auswertung	
	Evaluationsforschung und Interventionsforschung		4.2	Urteilen	154
3.1.2	Der Evaluator	103	4.2.1	Rangordnungen	155
	Soziale Kompetenz			Direkte Rangordnungen	
	Fachliche Kompetenz			Methode der sukzessiven Intervalle	
3.1.3	Rahmenbedingungen für Evaluationen	106		»Law of Categorical Judgement«	
	Wissenschaftliche und formale Kriterien		4.2.2	Dominanzpaarvergleiche	157
	Ethische Kriterien			Indirekte Rangordnungen	
3.2	Planungsfragen	109		»Law of Comparative Judgement«	
3.2.1	Hintergrundwissen	109		Die Konstanzmethode	
3.2.2	Wahl der Untersuchungsart	109		Das »Signalentdeckungsparadigma«	
	Evaluation durch Erkundung		4.2.3	Ähnlichkeitspaarvergleiche	170
	Evaluation durch Populationsbeschreibung			Die »klassische« multidimensionale Skalierung (MDS)	
	Evaluation durch Hypothesenprüfung			Die nonmetrische multidimensionale Skalierung (NMDS)	
				Die Analyse individueller Differenzen (INDSCAL)	

4.2.4	Ratingskalen	176	4.5.1	Alltagsbeobachtung und systematische Beobachtung	263
	Varianten für Ratingskalen			Kriterien der systematischen Beobachtung	
	Messtheoretische Probleme bei Ratingskalen			Modellierungsregeln	
	Urteilsfehler beim Einsatz von Ratingskalen		4.5.2	Formen der Beobachtung	266
	Mehrere Urteiler			Teilnehmende oder nichtteilnehmende Beobachtung?	
	Besondere Anwendungsformen von Ratingskalen			Offene oder verdeckte Beobachtung?	
4.2.5	Magnitude-Skalen	188		Nonreaktive Beobachtung	
4.3	Testen	189		Mehrere Beobachter	
4.3.1	Testethik	192		Apparative Beobachtung	
4.3.2	Aufgaben der Testtheorie	193		Automatische Beobachtung	
4.3.3	Klassische Testtheorie	193		Selbstbeobachtung	
	Die fünf Axiome der klassischen Testtheorie		4.5.3	Durchführung einer Beobachtungsstudie	269
	Die drei Testgütekriterien			Vorbereitung des Beobachtungsplanes	
	Die Multitrait-Multimethod-Methode (MTMM)			Ereignisstichprobe oder Zeitstichprobe?	
4.3.4	Item-Response-Theorie (IRT)	206		Technische Hilfsmittel	
	Itemcharakteristiken		4.5.4	Beobachtertraining	272
	Das dichotome logistische Modell			Beobachterübereinstimmung	
	Verallgemeinerungen und Anwendungen		4.6	Physiologische Messungen	278
	Latente Klassenanalyse		4.6.1	Methodische Grundlagen und Probleme	278
	Adaptives Testen			Allgemeine Messprinzipien	
	Klassische und probabilistische Testtheorie:			Messprobleme	
	Zusammenfassende Bewertung		4.6.2	Indikatoren des peripheren Nervensystems	280
4.3.5	Testitems	213		Kardiovaskuläre Aktivität	
	Itemformulierungen			Elektrodermale Aktivität	
	Ratekorrektur			Muskuläre Aktivität	
	Itemanalyse		4.6.3	Indikatoren des zentralen Nervensystems	286
4.3.6	Testskalen	221		Elektrophysiologische ZNS-Aktivität	
	Thurstone-Skala			Neurochemische Indikatoren	
	Likert-Skala			Bildgebende Verfahren	
	Guttman-Skala		4.6.4	Indikatoren endokriner Systeme	
	Edwards-Kilpatrick-Skala			und des Immunsystems	289
	Rasch-Skala			Aktivität endokriner Systeme	
	Coombs-Skala			Aktivität des Immunsystems	
4.3.7	Testverfälschung	231		Übungsaufgaben	
	Selbstdarstellung		5	Qualitative Methoden	295
	Soziale Erwünschtheit		5.1	Qualitative und quantitative Forschung	296
	Antworttendenzen		5.1.1	Qualitative und quantitative Daten	296
4.4	Befragen	236		Quantitative Daten	
4.4.1	Mündliche Befragung	237		Verbale Daten	
	Formen der mündlichen Befragung			Informationsgehalt	
	Aufbau eines Interviews			Vor- und Nachteile	
	Der Interviewer			Transformation qualitativer Daten	
	Die Befragungsperson			in quantitative Daten	
	Durchführung eines Interviews		5.1.2	Gegenüberstellung qualitativer	
4.4.2	Schriftliche Befragung	252		und quantitativer Verfahren	298
	Fragebogenkonstruktion			Nomothetisch versus idiografisch	
	Postalische Befragung			Labor versus Feld	
	Computervermittelte Befragung			Deduktiv versus induktiv	
	Delphi-Methode			Erklären versus Verstehen	
4.5	Beobachten	262			

5.1.3	Historische Entwicklung des qualitativen Ansatzes	302	6	Hypothesengewinnung und Theoriebildung	351
	Dominanz des quantitativen Ansatzes				
	Hermeneutik und Phänomenologie				
	Chicagoer Schule				
	Der Positivismusstreit				
	Qualitative Forschung als eigene Disziplin				
	Kanon qualitativer Methoden				
5.2	Qualitative Datenerhebungsmethoden	308	6.1	Theoriebildung im wissenschaftlichen Forschungsprozess	352
5.2.1	Qualitative Befragung	308	6.1.1	Exploration in Alltag und Wissenschaft	352
	Auswahlkriterien für qualitative Interviews			Exploration im Alltag	
	Arbeitsschritte bei qualitativen Interviews			Exploration in der Wissenschaft	
	Dokumentation einer Befragung		6.1.2	Exploration in Grundlagen- und Evaluationsforschung	354
	Techniken der Einzelbefragung		6.1.3	Inhaltliche und instrumentelle Voruntersuchungen	355
	Techniken der Gruppenbefragung		6.1.4	Exploration als Untersuchungstyp und Datenerhebungsverfahren	356
5.2.2	Qualitative Beobachtung	321	6.1.5	Vier Explorationsstrategien	357
	Beobachtung von Rollenspielen		6.2	Theoriebasierte Exploration	358
	Einzelfallbeobachtung		6.2.1	Theoriequellen	359
	Selbstbeobachtung			Alltagstheorien	
5.2.3	Nonreaktive Verfahren	325		Wissenschaftliche Theorien	
5.2.4	Gütekriterien qualitativer Datenerhebung	326	6.2.2	Theorieanalyse	360
	Objektivität			Zusammenfassung und Bewertung	
	Reliabilität			Vergleich und Integration	
	Validität			Formalisierung und Modellbildung	
5.3	Qualitative Auswertungsmethoden	328		Metatheorien	
5.3.1	Arbeitsschritte einer qualitativen Auswertung	329	6.2.3	Theoriebasierte Exploration: Zusammenfassung	364
5.3.2	Besondere Varianten der qualitativen Auswertung	331	6.3	Methodenbasierte Exploration	365
	Globalauswertung		6.3.1	Methoden als Forschungswerkzeuge	365
	Qualitative Inhaltsanalyse nach Mayring			Methodenvergleiche	
	Grounded Theory			Methodenvariation	
	Sprachwissenschaftliche Auswertungsmethoden		6.3.2	Methoden als Denkwerkzeuge	366
5.3.3	Gütekriterien qualitativer Datenanalyse	334		Analogien bilden	
	Gültigkeit von Interpretationen			Metaphern aufdecken	
	Generalisierbarkeit von Interpretationen		6.3.3	Methodenbasierte Exploration: Zusammenfassung	368
5.4	Besondere Forschungsansätze	336	6.4	Empirisch-quantitative Exploration	369
5.4.1	Feldforschung	336	6.4.1	Datenquellen	369
	Geschichte der Feldforschung			Nutzung vorhandener Daten	
	Arbeitsschritte in der Feldforschung			Datenbeschaffung durch Dritte	
5.4.2	Aktionsforschung	341		Eigene Datenbeschaffung	
	Methodische Grundsätze		6.4.2	Explorative quantitative Datenanalyse	371
	Praktische Durchführung			Einfache deskriptive Analysen	
5.4.3	Frauen- und Geschlechterforschung	343		Grafische Methoden: der EDA-Ansatz	
	Geschlecht als Konstrukt			Multivariate Explorationstechniken	
	Methodische Besonderheiten			Exploratives Signifikanztesten	
5.4.4	Biografieforschung	346		Data-Mining	
	Biografisches Material		6.5	Empirisch-qualitative Exploration	380
	Auswertungsverfahren				
	Genealogie				
	Psychohistorie				
	Übungsaufgaben				

6.5.1	Datenquellen	380	7.2.5	Der Bayes'sche Ansatz	455
	Nutzung vorhandener Daten			Skizze der Bayes'schen Argumentation	
	Datenbeschaffung durch Dritte			Diskrete Zufallsvariablen	
	Eigene Datenbeschaffung			Stetige Zufallsvariablen	
6.5.2	Explorative qualitative Datenanalyse	381		Schätzung von Populationsmittelwerten	
	Inventare			Schätzung von Populationsanteilen	
	Typen und Strukturen		7.2.6	Resamplingansatz	478
	Ursachen und Gründe		7.2.7	Übersicht populationsbeschreibender	
	Verläufe			Untersuchungen	479
	Systeme			Übungsaufgaben	
	Übungsaufgaben				
7	Populationsbeschreibende Untersuchungen	393	8	Hypothesenprüfende Untersuchungen	489
7.1	Stichprobe und Population	394	8.1	Grundprinzipien der statistischen	
7.1.1	Zufallsstichprobe	396		Hypothesenprüfung	491
	Zum Konzept »Repräsentativität«		8.1.1	Hypothesenarten	491
	Ziehung einer einfachen Zufallsstichprobe		8.1.2	Signifikanztests	494
	Probleme der Zufallsstichprobe			Zur Logik des Signifikanztests	
	Probabilistische und nichtprobabilistische			Ein Beispiel. Der t-Test	
	Stichproben		8.1.3	Probleme des Signifikanztests	498
7.1.2	Punktschätzungen	402	8.2	Varianten hypothesenprüfender	
	Zufallsexperimente und Zufallsvariablen			Untersuchungen	502
	Verteilung von Zufallsvariablen		8.2.1	Interne und externe Validität	502
	Kriterien für Punktschätzungen			Gefährdung der internen Validität	
	Parameterschätzung: Maximum-Likelihood-Methode			Gefährdung der externen Validität	
7.1.3	Intervallschätzungen	410	8.2.2	Übersicht formaler Forschungshypothesen	505
	Konfidenzintervall des arithmetischen Mittels		8.2.3	Zusammenhangshypothesen	506
	bei bekannter Varianz			Bivariate Zusammenhangshypothesen	
	Konfidenzintervall des arithmetischen Mittels			Multivariate Zusammenhangshypothesen	
	bei unbekannter Varianz			Kausale Zusammenhangshypothesen	
	Konfidenzintervall eines Populationsanteils			Zusammenfassende Bewertung	
7.1.4	Stichprobenumfänge	419	8.2.4	Unterschiedshypothesen	523
	Schätzung von Populationsanteilen			Kontrolltechniken	
	Schätzung von Populationsmittelwerten			Zweigruppenpläne	
7.1.5	Orientierungshilfen für die Schätzung			Mehrgruppenpläne	
	von Populationsstreuungen	423		Faktorielle Pläne	
7.2	Möglichkeiten der Präzisierung			Hierarchische Pläne	
	von Parameterschätzungen	424		Quadratische Pläne	
7.2.1	Geschichtete Stichprobe	425		Pläne mit Kontrollvariablen	
	Schätzung von Populationsmittelwerten			Multivariate Pläne	
	Schätzung von Populationsanteilen			Zusammenfassende Bewertung	
7.2.2	Klumpenstichprobe	435	8.2.5	Veränderungshypothesen	547
	Schätzung von Populationsmittelwerten			Experimentelle Untersuchungen	
	Schätzung von Populationsanteilen			Quasiexperimentelle Untersuchungen	
7.2.3	Die mehrstufige Stichprobe	440		Untersuchungspläne	
	Schätzung von Populationsmittelwerten			Veränderungshypothesen für Entwicklungen	
	Schätzung von Populationsanteilen			Veränderungshypothesen für Zeitreihen	
7.2.4	Wiederholte Stichprobenuntersuchungen	447		Zusammenfassende Bewertung	
	Schätzung von Populationsmittelwerten				
	Schätzung von Populationsanteilen				

8.2.6	Hypothesen in Einzelfalluntersuchungen	580	9.4.4	Abweichung eines Anteilswertes P von $p=0,5$	659
	Individuelle Veränderungen		9.4.5	Vergleich von zwei Anteilswerten P_A und P_B . . .	661
	Einzelfalldiagnostik		9.4.6	Häufigkeitsanalysen	661
	Zusammenfassende Bewertung			Kontingenztafel	
	Übungsaufgaben		9.4.7	Varianzanalysen	662
				Einfaktorielle Varianzanalyse	
9	Richtlinien für die inferenzstatistische Auswertung von Grundlagenforschung und Evaluationsforschung	599		Zweifaktorielle Varianzanalyse	
9.1	Statistische Signifikanz und praktische Bedeutsamkeit	602	9.4.8	Multiple Korrelation	668
9.1.1	Teststärke	602		Übungsaufgaben	
9.1.2	Theorie »optimaler« Stichprobenumfänge	604	10	Metaanalyse	671
9.2	Festlegung von Effektgrößen und Stichprobenumfängen	605	10.1	Zielsetzung	672
9.2.1	Effektgrößen der wichtigsten Signifikanztests	605	10.2	Auswahl der Untersuchungen	674
	Bedeutung der Effektgrößen		10.2.1	Selektionskriterien	674
	Klassifikation der Effektgrößen		10.2.2	Abhängige Untersuchungsergebnisse	675
9.2.2	Optimale Stichprobenumfänge für die wichtigsten Signifikanztests	627	10.3	Vereinheitlichung von Effektgrößen: das Δ-Maß	676
	Tabelle der optimalen Stichprobenumfänge		10.4	Zusammenfassende Analysen	681
	Erläuterungen und Ergänzungen		10.4.1	Homogenitätstest für verschiedene Δ -Maße . . .	681
	Verallgemeinerungen		10.4.2	Signifikanztest für den Gesamteffekt	681
9.3	Überprüfung von Minimum-Effekt- Nullhypothesen	635	10.4.3	Moderatorvariablen	682
9.3.1	Signifikanzschranken und Teststärkeanalysen	636	10.4.4	Teststärke von Metaanalysen	683
	Prüfung von H_{00}			Homogenitätstest	
	Prüfung von H_{01}			Signifikanztest	
	Prüfung von H_{05}			Moderatorvariablenanalyse	
	Hinweise zur Untersuchungsplanung		10.4.5	Ein kleines Beispiel	686
9.3.2	Transformation statistischer Test- und Kennwerte in die F-Statistik	643		Fünf Untersuchungen zum Lehrerurteil	
	Anwendungen		10.5	Probleme und Alternativen	693
	Zwei- und mehrfaktorielle Varianzanalysen		10.5.1	Signifikante und nichtsignifikante Untersuchungsergebnisse	695
	Kurzanleitung zur Nutzung von Tab. F11 (»Alles auf einen Blick«)		10.5.2	Exakte Irrtumswahrscheinlichkeiten	696
9.3.3	Zur Frage der »Bestätigung« von Nullhypothesen	650	10.5.3	Publikationsbias	697
	Beispiele für H_{00} -Hypothesen			Übungsaufgaben	
9.4	Beispiele für die Planung und Auswertung hypothesenprüfender Untersuchungen	655	Anhang		701
9.4.1	Vergleich von zwei Mittelwerten	656	Anhang A. Lösungen der Übungsaufgaben		702
	Unabhängige Stichproben		Anhang B. Glossar		723
	Abhängige Stichproben		Anhang C. Literatur- und Informationsquellen		747
9.4.2	Korrelation	658	Anhang D. Auswertungssoftware		751
9.4.3	Vergleich von zwei Korrelationen	659	Anhang E. Forschungsförderung		753

Anhang F. Tabellen	757	11 »Alles auf einen Blick«	804
1 Standardnormalverteilung	757	12 Untere Grenzen des 95%igen Konfidenzintervalls für ρ^2	822
2 Zufallszahlen	762	Anhang G. SAS-Syntax für die Berechnung einiger Konfidenzintervalle	827
3 t-Verteilungen	763	Literatur	829
4 Beta-Verteilungen (Abbildungen)	764	Namenverzeichnis	877
5 Beta-Verteilungen (Tabellen)	766	Sachverzeichnis	889
6 Iterationshäufigkeitstest	790		
7 Rangsummentest	794		
8 χ^2 -Verteilungen	799		
9 Fishers Z-Werte	802		
10 Arcus-sinus-Transformationen	803		