# Natur- und kulturwissenschaftliche Perspektiven in der Psychologie

## Wassilios Baros & Jürgen Rost (Hrsg.)

## Natur- und kulturwissenschaftliche Perspektiven in der Psychologie

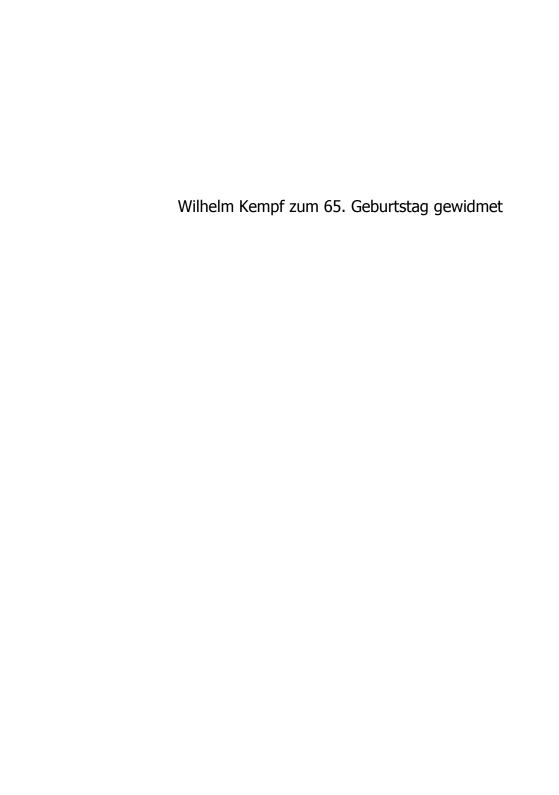
Methodologie – Methoden – Anwendungsbeispiele

© 2012 by verlag irena regener ostseestr. 109, d – 10409 berlin http://www.regener-online.de

Alle Rechte vorbehalten, auch die der fotomechanischen Wiedergabe und der Speicherung in elektronischen Medien.

Printed in Germany

ISBN 978-3-936014-28-0



## Inhalt

Zur Integration quantitativer und qualitativer Forschungsmethoden Wassilios Baros, Jürgen Rost & Hans Werbik
Friedens- und Konfliktforschung
Friedenspsychologie als zielorientierte Praxisforschung. Eine metatheoretischmethodologische Kritik Norbert Groeben & Johanna Ray Vollhardt
Elemente einer konstruktivistischen Konflikttheorie Heinz Messmer
Terrorismus-Berichterstattung und ihre Folgen Wolfgang Frindte, Nicole Haußecker & Jens Jirschitzka
"Worte, die wehtun, und Bilder, die schrei'n." Friedenskitsch am Beispiel von Songs im Eurovision Song Contest Susanne Jaeger & Michael Reimann
Mediendiskurs, Gewalt und gewaltfreie Konfliktbearbeitung. Das Journalism- in-the-New-World-Order-Projekt oder: von der Faszination einer Methode Stig-Arne Nohrstedt & Rune Ottosen
Methodologie und Methoden quantitativer Forschung
Nichtparametrische objektive Modelle für d-dimensionale Testdaten Hartmann Scheiblechner95
Ein marginaler Ansatz zur Schätzung von Item-Response-Modellen mit zufälligen individuellen Effekten  Martin Spiess
Das Beck Depressionsinventar (BDI-II). Psychometrische Analysen mit probabilistischen Testmodellen Ferdinand Keller
Qualitative und quantitative Methoden. Eine integrative Sichtweise  Jürgen Rost

# Subjektwissenschaftliche Perspektiven und integrative psychologische Forschungsmethodik

Qualitative versus quantitative Methoden: eine ewig währende Auseinandersetzung? Elfriede Billmann-Mahecha
Subjektwissenschaftliche Theorie und Empirie Morus Markard
Das argumentative Rückkoppelungsgespräch Wassilios Baros & Klaus-Dieter Reetz
Die Vision einer integrativen Psychologie im Spannungsfeld zwischen Natur- und Kulturwissenschaft Markus Kiefer & Klaus Hönig
Literatur
Anhang
Über die Autoren222
Index

## Zur Integration quantitativer und qualitativer Forschungsmethoden

Wassilios Baros, Jürgen Rost & Hans Werbik

Ein Buch zu konzipieren, welches der thematischen Bandbreite und der Relevanz des Werkes und Wirkens von Wilhelm Kempf in der Psychologie gerecht werden soll, ist aufgrund der Vielfalt seiner wissenschaftlichen Interessen- und Arbeitsfelder und der transdisziplinären Verankerung seiner Forschungsansätze gewiss kein leichtes Unterfangen. Die Beiträge dieses Buches können daher nur einige Facetten seines Werkes beleuchten, auch wenn sie einen recht weiten thematischen Bogen spannen: von Konflikt- und Friedensforschung, Medienanalysen und Friedensjournalismus über Antisemitismusforschung bis hin zu Testtheorie und Statistik sowie Methodologie und Methoden guantitativer und gualitativer Forschung. Diese wissenschaftlichen Tätigkeitsfelder sind nicht voneinander abgegrenzt, sondern vor dem Hintergrund eines konkreten wissenschaftstheoretischen bzw. methodologischen Verständnisses eng miteinander verknüpft, wobei das Zusammenspiel von Gegenstandsverständnis, Wissensideal und Aufgabenverständnis immer zu berücksichtigen ist. Wissenschaftstheorie, Methodologie und Methode als unterschiedliche Ebenen wissenschaftlichen Denkens und Handelns können in Kempfs Methodenlehre nicht separat voneinander beschrieben, sondern müssen in ihrer Brisanz, Funktion und unmittelbaren Dienlichkeit zur Genese, Analyse und Beantwortung relevanter Fragestellungen erörtert werden. Psychologische Forschungsmethoden sind immer im Zusammenhang mit den je konkreten Inhalten der Psychologie zu thematisieren. Dabei gilt es, der inhaltlichen Heterogenität der Psychologie mit einem konstruktiv-integrativen Blick zu begegnen, der es erst ermöglicht, die unterschiedlichen Wissensideale in ihrer Geltung anzuerkennen, welche ihrerseits je unterschiedliche methodische Schritte zur Geltungsprüfung von Theorien implizieren.

Davon ausgehend, dass sich quantitative und qualitative Forschungsmethoden nicht erst in den von ihnen favorisierten Verfahren unterscheiden, sondern bereits hinsichtlich ihres konkreten Gegenstandverständnisses und Wissensideals, wird im kempfschen Sinne für eine Methodenintegration plädiert, die weit über den Triangulationsgedanken eines naiven Methodenpluralismus ("Methodenmix") hinausgeht. Dabei ist eine grundlegende Orientierung an dem Primat des Gegenstandes vor der Methode zweckmäßig, um die Dichotomie zwischen quantitativen und qualitativen Methoden aufzuheben und deren Verhältnis zueinander neu zu definieren. Gleichwohl ist mit einer durch die Forschungsfragestellung begründeten Integra-

tionsentscheidung noch nichts darüber ausgesagt, ob und wenn ja, inwieweit diese Integration in ihrer konkreten methodischen Umsetzung auch tatsächlich gelingen kann. Dafür sind methodische Instrumente erforderlich, die in der Lage sind, Quantität und Qualität sinnvoll miteinander zu verbinden. Solch eine angemessene Kombination ermöglichen beispielsweise Latent-Class-Analysen in besonderer Weise, da sie auf die Klassifizierung von Antwortmustern in großen Stichproben abzielen und so Typisches aufspüren können, welches wiederum mit Hilfe qualitativer Methoden gegenstandsbezogen rekonstruiert werden kann. Kurz: Durch das Dynamisieren des Integrationsverständnisses im Methodendiskurs bestimmt Kempf das Verhältnis von Repräsentativität und Typikalität insofern neu, als (rekonstruierbare) Typikalität innerhalb der (erzielten) Repräsentativität sichtbar gemacht wird.

Wissenschaftstheoretisch geprägt durch die Tradition der konstruktiven Wissenschaftstheorie der Erlanger Schule entwickelt Kempf den Ansatz des methodologischen Konstruktivismus, welcher in Abgrenzung von diversen Spielarten des (sozialen) Konstruktivismus die Forderung nach transsubjektiver Geltung wissenschaftlicher Erkenntnisse hervorhebt und wissenschaftslogisch begründet. Auf dieser Grundlage wird der erforderliche methodologische Überbau geschaffen, durch welchen subjektwissenschaftliche Fragestellungen empirisch zugänglich gemacht und methodisch kontrolliert erforscht werden können.

Die Aufsätze des vorliegenden Buches gehören – nicht zufällig – sehr unterschiedlichen Bereichen an, markieren diverse "wissenschaftliche Lebensweisen", besetzen verschiedene Wirkungsfelder. Sowohl Textinterpretation und Inhaltsanalyse, die dem Wissensideal der qualitativen Forschung verpflichtet sind, als auch Psychometrie als die konsequenteste Umsetzung des quantifizierenden Wissensideals in der Psychologie sind in diesem Band vertreten. Wilhelm Kempf selbst hat mehr als Andere das wissenschaftliche Leben eines Wanderers zwischen zwei Welten gelebt, in denen er sich gleichermaßen zu Hause fühlt: Es geht, zum einen, um die Welt von Formeln und Zahlen. Und es geht, zum anderen, um die Welt, in der die Psychologie dazu herangezogen wird, politische Sachverhalte aufzuklären, die die Freiheit des Einzelnen in Frage stellen. Seine Arbeiten zu Konfliktkommunikation, zu Gewalt und Eskalation, zur subjektwissenschaftlichen Methodologie und Empirie bis hin zum methodologischen Konstruktivismus kennzeichnen nur einige seiner Schwerpunkte. Diese Schwerpunkte und Ansätze unterschiedlichster Provenienz wusste er immer auf eine beeindruckende, wissenschaftlich und terminologisch stringente Weise miteinander zu verknüpfen und Verbindungslinien für präzise Problemformulierungen mit gesellschaftskritischer Reichweite herzustellen. Wichtige Prinzipien in seiner Methodenlehre sind z. B. Anerkennung des Expertenstatus der Subjekte bezüglich ihrer eigenen subjektiven Wirklichkeit, Herausstellung des Prinzips der Kommunikation und Aufhebung der Trennung zwischen Öffentlichem und Privatem in der interpretativen Forschung sowie reflektierte Subjektivität des Forschers zur Ermöglichung zunehmender Transparenz des Analyseprozesses.

Im ersten Teil des Buches (Friedens- und Konfliktforschung) werden theoretische und methodologische Prämissen der Friedens- und Konfliktforschung expliziert, definitorische Abgrenzungen vorgenommen sowie Anwendungsbeispiele aus Forschungsprojekten zu friedenspsychologischen Fragestellungen präsentiert. Terminologische Vorentscheidungen sowie die Explikation des eigenen theoretischen und methodologischen Standpunktes sind wesentliche Prinzipien im wissenschaftlichen Denken Wilhelm Kempfs, der in Lehre und Forschung stets darum bemüht war, Grundbegriffe der Methodologie und Methodik mit konkreten substanzwissenschaftlichen Inhalten der Psychologie zu verbinden.

Norbert Groeben & Johanna Ray Vollhardt skizzieren am Beispiel der Friedenspsychologie die wissenschaftstheoretischen und methodologischen Struktur- und Prozessmerkmale einer zielorientierten Anwendungsforschung im Kontext und in Abgrenzung zu einer vor allem experimentell ausgerichteten Sozialpsychologie. Dabei geht es zunächst um den eigenständigen Wert einer auf externe Validität ausgerichteten Praxisforschung, die durch Theoriegeleitetheit (statt Methodenvorordnung) gekennzeichnet ist. Das impliziert nicht nur eine Theorienintegration, sondern vor allem auch eine konstruktive Gegenstands-Methodik-Interaktion, die ethischen Standards im Forschungsprozess eine konstitutive Rolle zuschreibt und eine Verbindung von quantitativen und qualitativen Methoden auf unterschiedlichen Analyseebenen (vom Individuum bis hin zu sozialen Institutionen etc.) einschließlich entsprechend komplexer statistischer Auswertungsmodelle realisiert – nicht zuletzt auch als Grundlage für die interdisziplinäre Rechtfertigung und Ausarbeitung des (positiven) Ziels Frieden.

Heinz Messmer befasst sich mit analytischen Grundlagen einer konstruktivistischen Konflikttheorie. Ausgangspunkt seiner Überlegungen ist ein systemtheoretisches Modell sozialer Konflikte, welches die konstitutiven Prozesse seiner Ausdifferenzierung Schritt für Schritt nachzuvollziehen erlaubt. Konflikte erweisen sich als sequenzielle Abfolgen der Widerspruchskommunikation, die einer typischen Strukturlogik folgen und die Bedingungen ihrer Fortführbarkeit bereits in sich tragen.

Wolfgang Frindte, Nicole Haußecker & Jens Jirschitzka untersuchen, wie in den Fernsehmedien über Terrorismus berichtet wird und wie diese Berichte wahrgenommen und interpretiert werden. Ihre Studie basiert zum einen auf der inhaltsanalytischen Auswertung von Hauptnachrichten im Fernsehen und zum anderen auf der Analyse von Interviews mit Rezipienten im Rahmen einer Panelstudie. Am Beispiel der Terrorismus-Berichterstattung lässt sich das Wechselspiel zwischen medialen Konstruktionen und individuellen Interpretationen aufzeigen. Die Medien- und Panelanalysen zeigen u. a., dass die privaten Fernsehsender stärker ausgeprägte Dramatisierungs- und Emotionalisierungsstrategien als die öffentlichrechtlichen nutzen. Auch bestätigt sich, dass sich Personen, welche überwiegend Privatfernsehen schauen, sowohl stärker durch den Terrorismus bedroht fühlen als auch stärker dazu tendieren, Muslime als Personen abzulehnen.

Der Beitrag von *Susanne Jaeger & Michael Reimann* reflektiert kritisch naive Vorstellungen von Frieden, wie sie sich im – in ihrer eigenen Terminologie – "Friedenskitsch" von Schlagertexten niederschlagen. Der Gedanke an "ein bißchen Frieden, ein bißchen Sonne" lasse zwar Individuen wenigstens vorübergehend an das Gute im Menschen und in der Welt glauben, habe jedoch die Entstehung pri-

vater Traumwelten zur Folge. Die Autoren bedienen sich der von Kempf für Nachrichtenmedien entwickelten Instrumente zur Analyse von Eskalations- und Deeskalationsorientierung und adaptieren sie auf Schlagertexte, die sich mit Krieg und/oder Frieden beschäftigen. Der Beitrag dokumentiert viele Textbeispiele aus über fünfzig Jahren Eurovision Song Contest und gipfelt in der Erarbeitung einer Checkliste, die es ermöglicht, das Phänomen Friedenskitsch in all seinen Facetten zu erfassen und zu analysieren.

Stig-Arne Nohrstedt & Rune Ottosen diskutieren das Spannungsverhältnis zwischen Mediendiskurs, Gewalt und gewaltfreier Konfliktbearbeitung anhand von Untersuchungsergebnissen des transdisziplinär und multilokal angelegten internationalen Medienforschungsprojekts Journalism in the New World Order. Auf der Basis der Analyse einer repräsentativen Anzahl von Pressemeldungen aus verschiedenenen Ländern wird der Frage nachgegangen, welche Rolle die Vision einer neuen Weltordnung in den Medienberichten spielte. Es wird überzeugend herausgearbeitet, wie es mittels Latent-Class-Analysen gelingt, den Einfluss von Propaganda – vermittelt durch subtile Formen und Stile der Berichterstattung – auf die Medienberichterstattung über den Golfkrieg 1991 gezielt und systematisch zu erforschen.

Die Beiträge des zweiten Teils (Methodologie und Methoden quantitativer Forschung) stehen insofern im Einklang mit Kempfs Denktradition, als hier psychometrische Methoden und Modelle sowie statistische Theorien psychologischer Daten vorgestellt werden, deren forschungspraktische Bedeutung weit über die Testpsychologie hinausreicht. Psychometrische Verfahren werden anwendungsbezogen präsentiert und die Diskussion über die klassische bzw. die probabilistische Testtheorie geht in Richtung einer Integration beider.

Hartmann Scheiblechner fokussiert die Frage, inwieweit die aus der Sicht der quantitativen Forschung notwendigen Eigenschaften der probabilistischen Testtheorie nicht auch in ordinalen Messstrukturen realisiert werden können. Speziell das Merkmal der spezifischen Objektivität ist nach Scheiblechner bereits "unterhalb" metrischer quantitativer Variablen definierbar. Sein Beitrag belegt eindrucksvoll, wie sich eine ordinale Testtheorie auf dem Axiom der spezifischen Objektivität aufbauen lässt. Anwendungsbeispiele zeigen die Berechnung von statistischen Kennwerten für die Modellgültigkeit und die Interpretation der Ergebnisse. Man kann den Beitrag sicherlich nicht als Brückenschlag zwischen quantitativen und qualitativen Methoden werten. Die Übertragung wesentlicher Eigenschaften metrischer Messsmodelle auf die Ebene ordinaler Messstrukturen ist jedoch ein Meilenstein auf dem Weg der Annäherung beider.

Martin Spiess schlägt einen einfachen Schätzer für die Itemschwierigkeiten in IRT-Modellen für binäre Antworten vor, der nicht von der korrekten Spezifikation der Korrelationsstruktur abhängt. Ist die inverse Responsefunktion analytisch verfügbar und ist die individuelle Designmatrix identisch für alle Personen und invertierbar, dann ist dieser Schätzer, der auf verallgemeinerten Schätzgleichungen (generalized estimating equations; GEE) basiert, analytisch verfügbar und asymptotisch effizient. In einer Simulationsstudie wird dieser GEE-Schätzer mit einem marginalen Maximum-Likelihood-(MML-)Schätzer verglichen, bei dem die zu be-

rechnenden Integrale mit Hilfe der adaptiven Gauss-Hermite-Quadratur approximiert werden. Aufgrund der Ergebnisse kann davon ausgegangen werden, dass der vorgeschlagene Schätzer in diesen Situationen dem MML-Schätzer vorzuziehen ist.

Ferdinand Keller gibt einen umfassenden Überblick über die historische Entwicklung des wohl berühmtesten Depressionsfragebogens (Beck Depression Inventory). Er diskutiert konkurrierende Testmodelle, die zur Analyse von BDI-Daten herangezogen werden können, und vergleicht die Eigenschaften des Fragebogens in qualitativ unterschiedlichen Stichproben miteinander. Damit lässt sich der Fragebogen über verschiedene Konstruktionen von Item- und Antwortformaten an die spezifischen Anforderungen einer diagnostischen Anwendung anpassen. Das Zusammenspiel und die wechselseitige Anpassung qualitativer und quantitativer Methodenelemente zeichnen diesen Text aus.

Dass qualitative und quantitative Methoden einander ergänzen, ist eine Position, die wohl in jedem Beitrag dieses Buches in der einen oder anderen Art und Weise vertreten wird. Der Aufsatz von *Jürgen Rost* ist insofern radikal, als er von einem erkenntnistheoretischen Modell ausgeht, das es erlaubt, den einzelnen Stadien des Erkenntnisprozesses unterschiedliche Methoden, also quantitative oder qualitative, zuzuordnen. Hierfür werden sechs Stadien des Forschungsprozesses unterschieden, deren Abfolge einen Zyklus des Spiralmodells definieren. Dabei kann es z. B. passieren, dass ein Forschungsvorhaben mit qualitativen Methoden der epistemischen Verortung des Forschungsvorhabens und der Hypothesengenerierung beginnt, dann aber quantitative Methoden der Messtheorie und der Inferenzstatistik anwendet. Wenn auch diese Art der Integration eher selten sein dürfte, so dient sie nicht nur der konzeptuellen Einordnung des eigenen Tuns, sondern vielleicht auch der Brauchbarkeit der empirischen Resultate.

Im dritten Teil des Buches (Subjektwissenschaftliche Perspektiven und integrative psychologische Forschungsmethodik) wird schwerpunktmäßig dessen zentrale Thematik behandelt, namentlich die Präzisierung einer Vision der Einheit in der Psychologie. Gefordert wird eine Synthese von Subjektivität und Objektivität, die zentralen Prinzipien einer vom Standpunkt des Subjekts ausgehenden psychologischen Forschung werden erläutert sowie Verfahren einer subjektpsychologischinterpretativen Forschung skizziert.

Elfriede Billmann-Mahecha gibt zunächst einen kurzen historischen Rückblick auf einige Kontroversen um das wissenschaftliche Selbstverständnis der Psychologie und stellt daran anschließend gängige dichotome Gegenüberstellungen qualitativer und quantitativer Ansätze grundsätzlich infrage, bevor die aktuellen Vorschläge, die unter dem Label Mixed Methods firmieren, hinsichtlich ihres Innovationsgehalts kritisch hinterfragt werden. Abschließend wird der mögliche komplementäre Erkenntnisgewinn qualitativer Forschung anhand einer Pilotstudie zur Selbstkonzeptforschung veranschaulicht und diskutiert.

*Morus Markard* stellt dar, wie sich die kritisch-psychologische Konzeption einer Forschung vom Standpunkt des Subjekts aus der Reflexion von Desideraten quantitativer wie qualitativer Herangehensweisen herausgebildet hat und warum der Begründungs- statt des Bedingtheitdiskurses Wissenschaftssprache der Psychologie sein sollte: Handeln ist nicht durch Bedingungen bewirkt, sondern in deren

subjektiver Bedeutung (Prämissen) begründet. Im Ergebnis ist die Sicherung wissenschaftlicher Objektivität nicht über die Herstellung und Kontrolle eines künstlich erzeugten zwischenmenschlichen Mangelzustandes zu gewährleisten, sondern über eine methodisch reflektierte soziale Selbstverständigung. Da sich auch in nomothetisch formulierten Theorien begründungstheoretische Annahmen aufzeigen lassen, lässt sich der Begründungsdiskurs nicht in eine hermeneutische Nische abschieben.

In der Tradition einer subjektpsychologisch-interpretativen Psychologie entwickeln *Wassilios Baros & Klaus-Dieter Reetz* in Anlehnung an den Argumentationsbegriff der konstruktiven Wissenschaftstheorie das Verfahren des argumentativen Rückkoppelungsgesprächs als forschungsmethodische Alternative zum konsensorientierten Ansatz der kommunikativen Validierung. Es zielt auf die Fortsetzung des hermeneutischen Zirkels ab, und zwar auf der Basis von Nachgesprächen, in denen Interpret und Interpretierte zu den Interpretationen argumentativ Stellung nehmen, ihre eigenen Wahrnehmungen und subjektiven Wirklichkeiten kommunizieren, intersubjektiv geteilte Wirklichkeiten 'aushandeln' und im Dialog ggf. auch Kriterien für die Überprüfung der Angemessenheit ihrer Handlungsprämissen erarbeiten.

Davon ausgehend, dass die wissenschaftliche Psychologie seit ihrer Gründung mit dem Gegensatz zwischen der Subjektbezogenheit ihres Forschungsgegenstandes und der Notwendigkeit wissenschaftlicher Objektivität konfrontiert ist, skizzieren *Markus Kiefer & Klaus Hönig* die Vision einer integrativen Psychologie, welche einen gangbaren Weg für die Bewältigung dieser Herausforderung aufzeigen könnte. Anhand von Beispielen aus der Gedächtnisforschung wird nachvollziehbar gemacht, wie ein naturwissenschaftlicher experimenteller bzw. neurowissenschaftlicher Ansatz, der auf die Analyse kognitiver Prozesse ausgerichtet ist, mit einem kulturwissenschaftlichen Ansatz, der auf die Rekonstruktion der subjektiven Erfahrungswelten zielt, kombiniert werden kann. Dabei werden beide Ansätze als komplementäre Forschungsparadigmen in einer integrativen Psychologie verstanden.

Alle Beiträge dieses Bandes beleuchten das Werk Wilhelm Kempfs aus unterschiedlichen Blickwinkeln und belegen gleichzeitig den nachhaltigen Einfluss seiner Arbeiten auf den aktuellen Diskurs in der Psychologie. Sie arbeiten den Begründungszusammenhang zwischen den Aufgaben der Wissensbildung, der logischen Struktur von Theorien und den Methoden ihrer Geltungsprüfung heraus und machen deutlich, dass methodische Fragen nur unter Berücksichtigung des Gegenstandsverständnisses, des Wissensideals und des Aufgabenverständnisses der jeweiligen *transdisziplinär* forschenden wissenschaftlichen Disziplin argumentationszugänglich sind und beantwortet werden können.

Last but not least: Das vorliegende Buchprojekt hätte in dieser Form nicht ohne die kompetente Unterstützung von Frau Dr. Irena Regener realisiert werden können. Nicht nur dafür, sondern auch für ihr persönliches Engagement bei der Konzeptualisierung dieser Idee und die engagierte Begleitung des Vorhabens von Anfang an bis hin zum Erscheinen des Buches möchten wir als Herausgeber ihr in besonderer Weise danken.



## Friedenspsychologie als zielorientierte Praxisforschung. Eine metatheoretisch-methodologische Kritik

Norbert Groeben & Johanna Ray Vollhardt

"Eine gegenstandsangemessene Psychologie ist nur in der Einheit von Naturund Kulturwissenschaft denkbar." Kempf 2009a, S. 37f.

"Frieden bedeutet also nicht einen Zustand konfliktfreier Harmonie, sondern einen bestimmten Modus der Konfliktbearbeitung." Kempf 1993, S. 54

#### 1 Fragestellung/en

Ein nicht unbeträchtlicher Teil der Friedenspsychologie basiert auf sozialpsychologischen Theorien oder ist sogar in der Sozialpsychologie angesiedelt (Boehnke et al. 2005; Cohrs & Boehnke 2008); diese zeigt allerdings immer noch und immer wieder ambivalente Reaktionen auf friedenspsychologische Forschung. Denn zum einen ist Frieden natürlich ein Wert, der gesellschaftlich positiv besetzt ist und daher auch für die Forschung ein zu realisierendes Ziel darstellt. Aber zum anderen führt gerade diese Wert- bzw. Zielorientierung der Friedenspsychologie zu Forschungsstrukturen, die den methodologischen Standards der traditionellen (experimentellen) Sozialpsychologie keinesfalls entsprechen (können und wollen). Die gesellschaftspolitisch unvermeidbare (explizite) Akzeptanz der Friedenspsychologie ist daher häufig mit fachpolitischen (impliziten) Vorbehalten verbunden. Wir wollen uns im Folgenden bemühen, diese Vorbehalte zu verringern, indem wir einige zentrale methodologische Charakteristika friedenspsychologischer Forschung herausarbeiten und metatheoretisch begründen - in der Hoffnung, dass die systematische Erklärung und Legitimation dieser spezifischen Forschungsstruktur deren Anerkennung in der Psychologie verbessern und damit zu einer Ausweitung friedenspsychologischer Forschungsbemühungen führen kann.

Die dafür essenziellen Fragestellungen sind im Prinzip schon in den zentralen Begriffen des Beitragstitels enthalten. Zunächst einmal impliziert der Gegenstand Frieden, wie erwähnt, ein konstitutives präskriptives Element (Sommer & Fuchs 2004). Friedenspsychologie ist sinnvoller- und notwendigerweise auf die Realisierung des angestrebten Ziels Frieden ausgerichtet, so dass sie das (webersche) Werturteilsfreiheitspostulat nicht übernehmen und vertreten kann (nähere Begründung vgl. Kap. 5). Zugleich bedeutet diese Zielorientierung, dass es sich nicht

(nur) um klassische Grundlagenforschung handelt, sondern dass ein (in traditioneller Terminologie) starker Anwendungsbezug vorliegt. Hinsichtlich der Unterscheidung von Grundlagen- und Anwendungsforschung hat die neuere wissenschaftstheoretische Diskussion allerdings zu der Einsicht geführt, dass das damit unterstellte Kaskadenmodell der Wissenschaft den Interdependenzen zwischen primär auf Erkenntnis und vor allem auf Verbesserung der praktischen Realität ausgerichteten Forschungsbemühungen nicht gerecht wird (vgl. Groeben 2006a). Wir sprechen daher mit Carrier (2004) lieber von epistemischer (traditionell: Grundlagen-) und praktischer (traditionell: Anwendungs-)Forschung. Es sind dann die Struktur- und Prozessmerkmale dieser (friedenspsychologischen) Praxisforschung herauszuarbeiten, und zwar in Form einer negativen wie positiven Kritik. Die negative Kritik besteht in der Elaboration der Unterschiede zur experimentellen Sozialpsychologie als (primär bzw. ausschließlich) epistemischer Forschung, die positive Kritik in der Rekonstruktion der spezifischen methodologischen Zielideen zur Realisierung der Friedens- als zielorientierter Praxisforschung. Das Vorgehen bei dieser Rekonstruktion ist ein metatheoretisch-methodologisches, insofern wir unter Rückgriff auf wissenschaftstheoretische Diskussionen möglichst konkrete methodologische Standards herausarbeiten wollen.

Diese Zielsetzung wollen wir erreichen, indem wir zunächst den eigenständigen Wert der Praxisforschung diskutieren (Kap. 2), wobei diese allerdings keineswegs als theorieabstinent zu verstehen ist, sondern im Gegenteil durch Theoriegeleitetheit (als Gegenpol zur Methodenvorordnung) gekennzeichnet ist (Kap. 3). Das impliziert zum einen eine (echte) konstruktive Gegenstands-Methodik-Interaktion (Kap. 4) und zum anderen eine interdisziplinäre Vernetzung, insbesondere in Bezug auf die Elaboration des Friedensziels (Kap. 5). Damit endet unsere metatheoretisch-methodologische Kritik bei einer grundlegenden Diskussion des Werturteilsfreiheitspostulats im interdisziplinären Kontext, aus der auch konstruktive Konsequenzen für die Weiterentwicklung der Sozialpsychologie auf theoretischer, methodologischer und praktischer Ebene gezogen werden können.

### 2 Zum eigenständigen Wert der praktischen (Friedens-)Forschung

Die Vorordnung und Höherwertung der (experimentell-)epistemischen (Grundlagen-) Forschung gegenüber der praktischen (Anwendungs-)Forschung geht nicht zuletzt auf das Dilemma zwischen interner und externer Validität zurück (Westermann 2000). Interne Validität wird durch eine (möglichst vollständige) Kontrolle aller Störvariablen erreicht, was allerdings eine Übertragung der Ergebnisse auf die von Störvariablen durchsetzte Alltagsrealität (externe Validität) kaum mehr möglich macht. Interne Validität führt also zu Ergebnissen, bei denen zwar der Status als Erkenntnis (relativ) sicher, deren Brauchbarkeit für die Alltagsrealität aber (absolut) unsicher ist. Konzentriert man sich jedoch von vornherein auf die externe Validität, wird wegen der vielen Störvariablen bereits der Erkenntnisstatus der Ergebnisse unsicher, was die Frage nach der Brauchbarkeit sinnlos bzw. zumindest überflüssig macht. Aus Sicht der klassischen Methodenlehre liegt hier also ein

(asymmetrisches) Dilemma zwischen zwei Übeln vor, bei dem die Fokussierung auf die interne Validität allerdings als das kleinere angesehen wird! (Campbell & Stanley: vgl. Bortz & Döring 2002, S. 56f., S. 504f.) Das rechtfertigt die Vorordnung und Höherwertung der epistemischen gegenüber der praktischen Forschung.

Die wissenschaftstheoretische und -historische Diskussion hat jedoch gezeigt, dass diese hierarchische Sequenzierung der realen Entwicklung von Theorien nicht gerecht wird (Carrier 2004; Groeben 2006a). Vielmehr entstehen wissenschaftliche Hypothesen und Forschungsprogramme sowohl aus konzeptuell-erkenntnisgeleiteten Frageperspektiven als auch aus dem praktischen Problemdruck der Alltagsrealität. In der Allgemeinen Psychologie stellt das Multi-Tasking-Paradigma ein typisches Beispiel dar, in der (friedenspsychologischen) Sozialpsychologie wurde durch die Schrecken des 2. Weltkrieges die Aggressions- und Autoritarismusforschung angeregt (vgl. z. B. Jones 1998) sowie zu Zeiten des Kalten Krieges die Forschung zu Feindbildern und zur Abschreckungstheorie (vgl. Christie et al. 2008). In der wissenschaftlichen Theorieentwicklung ergänzen sich solche epistemischen und praktischen Forschungsanstrengungen und befruchten sich gegenseitig. Es ist daher eher ein Kontinuum von Übergängen bzw. Kombinationen des epistemischen und praktischen Erkenntnisinteresses anzusetzen. Daraus folgt allerdings, dass die methodologische Prävalenz der epistemischen Forschung mit dem Königsweg der (labor-)experimentellen Methodik (begründet mit der Sicherung der internen Validität) aufzugeben ist.

Das betrifft vor allem auch die unter interner Validitätsperspektive zentrale Bedingungsvariation auf Seiten der unabhängigen Variablen (UV). Epistemisch-experimentelle Forschung ist - entsprechend dem naturwissenschaftlichen Vorbild der Klassischen Physik – dadurch charakterisiert, dass die (potenziellen) UVen möglichst molekular gegeneinander variiert werden, damit man mit größtmöglicher Sicherheit sagen kann, welcher Effekt auf Seiten der abhängigen Variablen (AV) auf welche Bedingung (auf Seiten der UV) zurückzuführen ist. Eine solche molekulare Ausdifferenzierung und Gegeneinandervariation der UV-Seite ist nun aber bei praktischer Forschung z. T. gar nicht möglich, auf jeden Fall häufig nicht sinnvoll. Wenn in der Alltagsrealität nur bestimmte (molare) Bedingungscluster vorkommen, ist es weitgehend sinnlos, zu versuchen, diese Cluster unter der überzogenen Perspektive der internen Validität aufzuspalten, um nach Einzeleffekten und deren Erklärungskraft (gleich Varianzaufklärung) zu suchen (Bungard et al. 1988). Einen paradigmatischen Fall stellt hier die Relation von positiven und negativen (Einstellungs-)Variablen dar, z. B. Rachebedürfnis und Perspektivenübernahme; deren Gegeneinandervariation ist unsinnig, weil sie sich psychologisch eben gegenseitig ausschließen. Aber auch bei (von der Wertung her einsinnigen negativen oder positiven) Variablenclustern (wie z. B. Versöhnungsbereitschaft) macht es wenig Sinn, die Einzelkomponenten (z. B. Perspektivenübernahme, solidarisches Mitleid, Akzeptanz übergreifender humaner Werte etc.) aufzuspalten und experimentell zu variieren; angezeigt sind höchstens Längsschnittstudien zur Aufklärung darüber, in welcher Reihenfolge der Komponenten sich am besten ein solches Cluster (wie Versöhnungsbereitschaft) entwickelt.

Es gibt also epistemische Fragen, die in der Alltagsrealität nicht als Problem vorkommen und daher eine dysfunktionale Überdifferenzierung darstellen. Deshalb kann und sollte praktische Forschung legitimerweise von prototypischen Bedingungsclustern bzw. -kombinationen ausgehen – auch um eine zureichende interne Validität der Erklärungshypothesen zu erreichen! Ein hervorragendes Beispiel dafür ist die Studie von Kempf (2010), in der unter Anwendung der Latent-Class-Analyse die typischen Strukturen antisemitischer Einstellungen herausgearbeitet wurden.

Das bedeutet allerdings nicht, dass in praktischer Forschung das Verfahren des Gegeneinandervariierens gänzlich irrelevant ist. Es bezieht sich hier lediglich (primär) auf die Relation der (theoretisch) zentralen Bedingungscluster und (potenziellen) Störvariablen. Praktische Forschung muss gerade versuchen, die Effektivität der behaupteten UV-Cluster auf die angenommenen AV auch im (durch Störvariablen "verunreinigten") Feld nachzuweisen. Dazu sind dann (gerade auch) nicht-experimentelle Versuchspläne einzusetzen; von quasi-experimentellen Designs über Umfrage- und Feldstudien bis zu qualitativ-quantitativ integrativen Versuchsanordnungen (Hussy et al. 2010, Schreier 2006a, b). Im Prinzip kann man von dieser Perspektive aus die Hierarchierelation zwischen epistemischer und praktischer Forschung sogar genau entgegengesetzt zur bisherigen (üblichen) Sichtweise rekonstruieren, sozusagen vom Kopf auf die Füße stellen. Denn in labor-experimenteller (epistemischer) Forschung lassen sich wegen der Kontrolle der Störvariablen selbst vergleichsweise wenig robuste Hypothesen bewähren. Die Nagelprobe kommt erst, wenn sich die empirische Geltung gegen die Belastung durch verschiedenste hinderliche Variablen(-einflüsse) erweisen muss. Extrem formuliert: Man kann die laborexperimentelle epistemische Forschung als eine Art heuristisches Stadium für die Prüfung aussagekräftiger empirischer Gesetzmäßigkeiten ansehen; denn praktische Aussagekraft kann berechtigterweise erst dann behauptet werden, wenn die Bewährung in der komplexen Gemengelage von Wirkungen und Störeinflüssen der Alltagsrealität gelungen ist, also externe Validität vorliegt. Friedensforschung kann damit als paradigmatisches Feld für die Zusammenführung von Erkenntnisinteressen und Praxisrelevanz in einem interdisziplinären, methodologisch integrativen Gegenstandsverständnis fungieren.

#### 3 Theoriegeleitetheit statt Methodenvorordnung

Zur generellen metatheoretisch-methodologischen Charakterisierung von psychologischen Forschungsprogrammen unterscheiden Brinberg und McGrath zwischen Konzept-(Theorie-), Methoden- und Gegenstandsgeleitetheit der Forschung. Als optimales Ablaufmuster setzen sie die Sequenz KGM an (McGrath & Brinberg 1984). Interessanterweise kommt die praktische Friedensforschung dieser Abfolge sehr viel näher als die epistemisch-experimentelle Sozialpsychologie, für die eher das Ablaufmuster MKG gilt. Denn die Festlegung auf das Experiment als Königsweg der Erkenntnis stellt eine Methodenvorordnung dar, die nicht nur die erforschbaren Gegenstände beschränkt, sondern auch negative Konsequenzen für den konzep-

tuellen Bereich der Theorieentwicklung zeitigt. Entsprechend ist, im Vergleich zur Mainstream-Sozialpsychologie, die sozialpsychologische Friedensforschung durch eine größere Methodenvielfalt gekennzeichnet (Christie 2006; Vollhardt & Bilali 2008).

Die negativen Konsequenzen für die Theorieentwicklung gehen ebenfalls von der angesprochenen Prävalenz der internen Validität mit dem Prinzip des Gegeneinandervariierens von UVen (und der Kontrolle von Störvariablen) aus. Denn durch die Ausdifferenzierung der UVen werden auf der Ebene der Theorieentwicklung die Wenn-Komponenten der Hypothesen immer weiter spezifiziert. Es werden Moderator- und Mediatorvariablen ausdifferenziert, die zu einer Erhöhung der empirischen Geltung der jeweiligen Hypothese/Gesetzmäßigkeit führen (sollen). Dieser Validitätsgewinn soll erreicht werden, indem der Informationsgehalt der Wenn-Komponente (der Hypothese/Gesetzmäßigkeit) erhöht wird. Allerdings bedeutet eine solche Erhöhung des Informationsgehalts der Wenn-Komponente bei gleichbleibender Dann-Komponente (AV-Seite) ein Absinken des Informationsgehalts des Gesamtsatzes (vgl. schon Opp 1976). Das lässt sich logisch nachweisen, ist aber auch intuitiv völlig einsichtig, weil durch die Ausdifferenzierung der Wenn-Komponente (UVen) eine immer spezifischere Bedingungskonstellation herausgearbeitet wird, die als Verursachung eines bestimmten Wirkeffekts (AV) gelten kann; es werden also immer mehr (anders geartete) Bedingungskonstellationen als Ursachen dieses Effekts ausgeschlossen. Eine Erhöhung des Informationsgehalts der gesamten Hypothese ist nur möglich, wenn der Informationsgehalt der Dann-Komponente gesteigert wird, wenn also neue AVen hinzukommen, die mithilfe der theoretisch postulierten Bedingungskonstellationen (ebenfalls) erklärt werden können.

In diesem Spannungsfeld von Informationsgehalt auf Seiten der Wenn- vs. Dann-Komponente von Hypothesen/Gesetzmäßigkeiten erfüllen, wie leicht ersichtlich, epistemische (Grundlagen-) und praktische (Anwendungs-)Forschung komplementär divergierende Funktionen. Die (epistemisch bedingte) Ausdifferenzierung der Wenn-Komponente von Theorie-Hypothesen führt (unvermeidlich) zu einer Parzellierung der Theorienlandschaft mit einer (wachsenden) Vielzahl von Minitheorien (Groeben 2003). Denn zum einen wird der Geltungsbereich der Ausgangshypothese/-theorie durch die präzisierende Ausarbeitung der Variablenkonstellation/en auf UV-Seite zunehmend eingeschränkt; d. h. die Theorie wird vom Geltungsbereich her beschränkter, kleiner (eben "mini"). Zum anderen werden dadurch auch neue, weitere Theorien möglich, die sich mit dem frei gewordenen Problemraum der nicht mehr (in die präzisierte Theorie) einbezogenen Bedingungskonstellationen beschäftigen. Das bedeutet, dass permanent neue Theorieansätze entstehen (müssen), die ähnlich sind und trotzdem Eigenständigkeit behaupten, insofern sie ihre eigenen Bedingungskonstellationen in paradigmatischen Untersuchungen (examplars sensu Kuhn 1967) festschreiben. Da sie im Folgenden auch wieder den Weg weiterer Spezifizierung und damit "Minimierung" gehen, ist eine außerordentlich parzellierte Theorienlandschaft mit einer Vielzahl von (ähnlichen, aber Unähnlichkeit, Eigenständigkeit behauptenden) Minitheorien das Resultat. Das gilt nicht zuletzt auch für die experimentelle Sozialpsychologie, in der z. B. der Theorieansatz des fundamentalen Attributionsfehlers ausdifferenziert worden ist zum ultimativen Attributionsfehler, zum Akteur-Beobachter-Effekt der Attribution und der selbstwertdienlichen Verzerrung (Aronson et al. 2008). Als fundamentaler Attributionsfehler wird die Tendenz bezeichnet, dass man bei der Zuschreibung von Handlungsursachen die Persönlichkeitsfaktoren in Relation zu den Situationsbedingungen überschätzt. Für den Fall, dass die Persönlichkeitsfaktoren von der Zugehörigkeit zu einer sozialen Gruppe bestimmt werden, ist die Konzeptualisierung als ultimativer Attributionsfehler vorgeschlagen worden, während sich das Konstrukt des Akteur-Beobachter-Effekts auf das Phänomen bezieht, dass diese Attributionsfehler nur von der Beobachterposition aus gemacht werden, weil man als Akteur die situativen Bedingungen kennt und adäquat(er) berücksichtigt. Die selbstwertdienliche Verzerrung schließlich bezieht sich auf Fehlattributionen, die durch das Streben nach einem positiven Selbstwertgefühl zustande kommen (ebd., S. 108ff.)

Dieser Parzellierungsdynamik setzt die praktische (auch gerade Friedens-)Forschung eine komplementäre Integrationsdynamik entgegen (vgl. z. B. Vallacher et al. 2010). Das Bemühen um extern valide, in der Alltagsrealität brauchbare Theorien enthält im- und explizit drei Wege zu einer solchen Integrationsdynamik. Zum einen können dadurch unterschiedliche Konstruktbenennungen, die sich aber auf vergleichbare Variablen beziehen, in übergreifenden Konstruktexplikationen zusammengefasst werden; auf diese Weise sind sozusagen Theoriefamilien rekonstruierbar, die einen erheblichen Integrationseffekt bieten können. Zum zweiten werden für die unterschiedlichen (epistemischen) Theorien zentrale Geltungsbereiche identifiziert, die in einem übergeordneten Theoriemodell integriert werden, indem durch die Kombination von Geltungsschwerpunkten der Einzeltheorien ein umfassend(er)er Geltungsbereich abgedeckt wird. Und das bedeutet zum dritten, dass dadurch der Informationsgehalt der Dann-Komponente von Hypothesen und damit der übergeordneten Theoriemodelle erweitert wird (indem neue AVen in den Erklärungsbereich einbezogen werden).

Ein Beispiel, das alle drei Integrationsperspektiven in sich vereint, bietet die Selbstwahrnehmung von Gewaltopfern. Hier kann man zwischen einer Konzeptualisierung als spezifischer Opfergruppe, in der die Unterschiede zu anderen Opfern hervorgehoben werden, und einer generelleren Konzeptualisierung differenzieren, in der die Ähnlichkeiten zu anderen Opfergruppen im Vordergrund stehen. Die Theorie der Konzeptualisierungsebenen (level of construal) bietet für diese (und ähnliche Unterscheidungen) den integrativen Rahmen (Trope & Liberman 2010), weil die höhere Konzeptualisierungsebene zu inklusiven Opferkategorisierungen führt, die niedrigere Ebene zu spezifischen, exklusiven Opferkategorien (Vollhardt 2011). Diese Rahmentheorie ermöglicht problemlos die Einbeziehung eines wichtigen empirischen Ergebnisses, nämlich dass eine übergreifende Kategorisierung von Opfergruppen zwar grundsätzlich positive, z. B. solidarische Handlungen gegenüber anderen Opfern auslöst, u. U. aber auch zu Reaktanz führen kann, wenn eine Opfergruppe das Gefühl hat, in ihrem spezifischen (Opfer-)Schicksal nicht genügend

von der Gesellschaft anerkannt zu sein (Vollhardt 2011; vgl. auch Hornsey & Hogg 2000). Die dadurch bedingten negativen Reaktionen müssen mit den positiven Reaktionstendenzen der inklusiven Opferkategorisierung zusammengeführt werden und erschließen im Optimalfall weitere (konstruktive) Reaktionsmöglichkeiten, nämlich prosoziale Handlungen gegenüber allen möglichen Minoritäten, benachteiligten Gruppen etc. – vorausgesetzt, das eigene Opferschicksal wird innerhalb einer inklusiven Opferkategorisierung angemessen anerkannt (Vollhardt 2009b).

Die Prävalenz der externen Validität in der praktischen Friedensforschung könnte also im Optimalfall zu dem quasi-paradoxalen Effekt führen, dass in ihr im Vergleich zur epistemisch-experimentell ausgerichteten Sozialpsychologie (mit ihrer Methodenvorordnung) eine stärkere Theoriegeleitetheit und vor allem Theorienintegration erreicht wird (vgl. Coleman 2006; Vallacher et al. 2010), von der die epistemische Sozialpsychologie auf Dauer durchaus profitieren kann.

#### 4 Konstruktive Gegenstands-Methodik-Interaktion

Es ist eine metatheoretische Binsenweisheit, dass es eine Passung von Gegenstand und Methode geben muss; mit dem Geologenhämmerchen lassen sich nun einmal keine psychischen Prozesse untersuchen (Groeben & Westmeyer 1981; Hussy et al. 2010). Methodologisch resultiert daraus die Forderung nach einer (möglichst gleichgewichtigen) Interaktion von Gegenstand und Methode; das heißt: Weder soll die Methode den Gegenstand dominieren und damit die Gegenstandsadäquatheit beschränken, noch darf die angestrebte Gegenstandsadäquatheit die Systematik des methodischen Zugriffs einschränken oder aufheben. Zwar wird diese methodologische Forderung durchweg vertreten, aber keineswegs immer erfüllt. Das gilt nicht zuletzt für die epistemisch-experimentelle Sozialpsychologie, deren Ablaufmuster MKG eine (beschränkende) Anpassung des Gegenstandes an die Methodik impliziert. Dagegen kommt das Ablaufmuster KGM der praktischen Friedensforschung dieser Forderung nach einer konstruktiven, gleichgewichtigen Gegenstands-Methodik-Interaktion sehr viel näher. Die Passung von Gegenstand und Methodik manifestiert sich (zumindest) auf vier Ebenen:

1. Dabei geht es zunächst einmal um die Rolle der Ethik in der Forschung. In der Friedenspsychologie führen die konstitutiven (positiven) anthropologischen Zielsetzungen unter der Perspektive einer Kohärenz von Ziel und Weg (der Forschung) zu besonders hohen ethischen Anforderungen, d. h. es müssen alle (methodischen) Vorgehensweisen vermieden werden, die Beeinträchtigungen oder gar (psychische) Verletzungen von Untersuchungsteilnehmer(inne)n implizieren oder auch nur riskieren (vgl. Schuler 1980; Kimmel 1996). Das bedeutet, dass eine große Anzahl von negativen Bedingungen, deren Wirkung und Überwindung Friedenspsychologie zum Thema hat, aus ethischen Gründen nicht (experimentell) manipuliert werden darf! Dies betrifft z. B. das Erleiden direkter ethnopolitischer Gewalt genauso wie Erlebnisse struktureller Gewalt (Staub & Vollhardt 2008; Vollhardt 2009a): Das Spektrum reicht von der Frage, ob sich Erniedrigung auf die Unter-

stützung von Selbstmordattentaten auswirkt (Ginges & Atran 2008) über Diskrimierung und Rassimus (Jasinskaja-Lahti et al. 2006) sowie Antisemitismus (Kempf 2010) bis zu Versuchsmaterialien, die autoritäre oder kriegsjournalistische Botschaften enthalten (vgl. Kempf 1994a) etc. Stattdessen verfolgt friedenspsychologische Forschung oft Fragestellungen und Forschungsdesigns, die die Verringerung struktureller Gewalt zum Ziel haben (vgl. Vollhardt & Biliali 2008) bzw. auf emanzipatorisches Potenzial und die Förderung von Frieden ausgerichtet sind (vgl. Martin-Baro 1994). Diese Vorordnung der Ethik in der Forschung stellt die entscheidende Rechtfertigung für die schon angeführte Berücksichtigung von nichtexperimentellen Untersuchungsdesigns (Feldstudien, Umfragestudien, qualitative Versuchspläne etc.) dar.

- 2. Das darin zum Ausdruck kommende (Qualitäts-)Merkmal der Integration von quantitativen und qualitativen Ansätzen manifestiert sich aber nicht nur im Bereich der Versuchsplanung, sondern auch der Erhebungsmethoden. Praktische friedenspsychologische Forschung muss notwendigerweise insbesondere das psychische Erleben der Menschen erheben, das für die Überwindung von Gewalt und die Sicherung von Frieden (mit-)entscheidend ist. Und auch dafür gilt, dass quantitative und qualitative Verfahren in einer gegenstandsangemessenen, konstruktiven Kombination eingesetzt werden sollten (Hussy et al. 2010, Kempf 2003b, 2008; Kempf et al. 2000, Schreier 2006c). So wurden beispielsweise in einer Evaluationsstudie zu friedensfördernden Radioprogrammen in Rwanda nicht nur Fragebogenverfahren eingesetzt, sondern auch Gruppendiskussionen und Entscheidungen über die Verteilung einer knappen Ressource beobachtet, um die Wirksamkeit der Intervention in Hinblick auf Kooperation und friedfertiges Verhalten zu überprüfen (Paluck 2009).
- 3. Diese Einbeziehung von qualitativen Erhebungsmethoden bedeutet allerdings keineswegs einen (unzulässigen) Individualismus. Um als praktische Forschung erfolgreich sein zu können, muss die Friedenspsychologie unbedingt den sozialen Kontext einbeziehen, und zwar sowohl auf der Ebene der sozialen Interaktionen als auch der sozialen Institutionen. Friedenspsychologische Forschung muss also Daten auf Mikro-, Meso- und Makroebene (von Individuen über Gruppen/Institutionen bis zu Gesellschaften/Nationen) berücksichtigen (Christie et al. 2008, Vollhardt & Bilali 2008) und dadurch eine Integration von methodologischem Individualismus und Holismus/Kollektivismus leisten (Groeben 1999). Auch diese Integrationsdynamik kann eine fruchtbare Rückkoppelung für das umfassendere disziplinäre Teilgebiet der (epistemischen) Sozialpsychologie(-forschung) darstellen und neue Erkenntnisse aufdecken. Zum Beispiel konnte Elcheroth (2006) in einer Mehrebenenanalyse aufzeigen, dass Kriegsopfer auf der individuellen Ebene Menschenrechte weniger unterstützten als Nicht-Opfer, während auf der kollektiven Ebene in Gemeinden mit einer größeren Anzahl von Kriegsopfern mehr Unterstützung für Menschenrechte zu finden war.
- 4. Und nicht zuletzt gehört zu dieser methodologischen Integration eine Ausweitung von Erklärungskonzepten sowie statistischen Auswertungsmodellen. Das Experiment als Königsweg der epistemischen Forschung ist auf die Sicherung von

Kausalität nach dem Vorbild der Naturwissenschaften ausgelegt. Das bezieht sich auf die strukturelle Erklärung als Antwort auf Warum-Fragen (Subsumptionsmodell der Erklärung). Praktische Forschung kann und muss aber darüber hinaus auch prozessorientierte Wie-Erklärungen einbeziehen, um die Genese und Realisierung angestrebter Zielzustände (hier: Frieden) umfassend aufzuklären. Das impliziert den Einsatz entsprechend komplexer, sophistizierter statistischer Auswertungsmodelle. Entgegen dem üblichen Stereotyp kann sich nämlich die experimentell-epistemische Forschung auf Standardauswertungen nach dem Allgemeinen Linearen Modell beschränken, während die Integration von gualitativen und guantitativen Versuchsplänen sowie Erhebungsmethoden und Erklärungsmodellen Auswertungsmodelle verlangt, die Ergebnismuster ienseits einer punktlinearen Kausalität zu sichern vermögen (von Mehrebenen- über Zeitreihen- bis zur Funktionalen Datenanalyse, Data mining etc.; vgl. Groeben 2006b; Beispiele aus der Friedenspsychologie in Bezug auf die Mehrebenenanalyse: Elcheroth 2006; die Pfadanalyse: Tausch et al. 2009; Strukturgleichungsmodelle mit multiplen Mediatoren: Canetti-Nisim et al. 2009; Latent-Class-Analyse: Kempf 2010; sowie Computersimulationen von nonlinearen dynamischen Modellen: Vallacher et al. 2010). Auch dies sind Aspekte, die zu einem fruchtbaren Austausch zwischen epistemisch-sozialpsychologischer und praktisch-friedenspsychologischer Forschung führen können und sollten. Dieser Austausch eröffnet dann ganz grundsätzlich und umfassend die Möglichkeit einer integrativen Zusammenführung von methodologischen Traditionen, die aufgrund der üblichen Entgegensetzung von interner und externer Validität, von Grundlagen- und Anwendungsforschung unnötigerweise als sich weitgehend ausschließend angesehen werden. Das betrifft z. B. bereits die grundlegende Unterscheidung von objekt- und subjektwissenschaftlichem Zugang der Forschung. Eine umfassende sozial-/friedenspsychologische Praxisforschung muss selbstverständlich objektive (Rahmen-)Bedingungen auf überindividuellen Analyseebenen einbeziehen und mit den subjektiven Sicht- und Erlebensweisen der Individuen verbinden. Das impliziert, dass – wo immer es möglich ist – qualitative Erhebungsverfahren ebenso in experimentellen Versuchssettings einzusetzen sind wie experimentelle Designs bei der Erforschung subjektiver Verarbeitungshorizonte. Letztlich löst sich dadurch sogar die klassische Verstehen-Erklären-Kontroverse (vgl. Apel 1979) auf (wie es auch von der neueren wissenschaftstheoretischen Diskussion erarbeitet worden ist: vgl. Schurz 1990): Eine zureichende Erklärung friedensrelevanter Prozesse ist nur durch das Verstehen der dabei essenziellen Erlebenshorizonte möglich, und das Verstehen der Verarbeitungsprozesse kann nur in Verbindung mit validen Erklärungen praktisch (friedenssichernd) relevant werden!

#### 5 Interdisziplinäre Zielelaboration

Einer der auffälligsten Unterschiede zwischen der epistemisch-experimentellen Sozial- und der praktischen Friedenpsychologie dürfte die Absage der Friedenspsychologie an das webersche Werturteilsfreiheitspostulat (WFP: Albert 1968) sein. Termini zur Bezeichnung von Gewalt wie "ethnische Säuberung", Folter, Vertreibung, Genozid etc. sind nun einmal wertungshaltig bzw. richtiger: Sie lassen sich

nicht rein deskriptiv gebrauchen, ohne die damit bezeichneten Prozesse implizit aufzuwerten. Deshalb ist die entscheidende Frage nicht, ob man Werturteile als wissenschaftliche Aussagen zulässt, sondern wie man deren Wissenschaftlichkeit durch rationale Begründung sichern kann (Erb 1997, Groeben 1999). Denn Weber hat das WFP vor allem deshalb aufgestellt, weil präskriptive Sätze einen größeren (eben den wertenden) Bedeutungsumfang als deskriptive Sätze besitzen und daher aus letzteren nicht (logisch zwingend) abgeleitet werden können. Allerdings ist logisches Schlussfolgern nicht die einzige Möglichkeit einer rationalen Diskussion und Rechtfertigung (vgl. Kempf 1978). Und es steht z. B. mit der Ziel-Mittel-Argumentation (heute) durchaus ein Instrument zur Verfügung, das eine rationale Kritik und Begründung auch von Werturteilen ermöglicht. Dabei werden deskriptive und präskriptive Sätze so zueinander in Beziehung gesetzt, dass von den erwünschten oder unerwünschten Wirkungen aus die positive oder negative Wertung der die Wirkung auslösenden Bedingungen begründet wird. Eine empirische Falsifikation der deskriptiven Wirkungshypothese dekuvriert die Wertung also als unbegründet (negative Kritik), eine Bewährung der empirischen Hypothese stellt eine (relative) Legitimation der übergeordneten Wertprämissen dar (König 1975, Groeben 1986). Um es an dem oben angeführten Beispiel der Opferkategorisierung zu verdeutlichen: Wenn eine inklusive Opferkategorisierung ohne Anerkennung des spezifischen Schicksals einer bestimmten Opfergruppe nicht zu der angenommenen Wirkung prosozialer Einstellungen gegenüber anderen Opfergruppen führt, dann ist damit eine positive Bewertung der inklusiven Opferkategorisierung als unbegründet abzulehnen. Die Anerkennung des spezifischen Opferschicksals ist dagegen positiv zu bewerten, sofern sie zu den behaupteten prosozialen Effekten führt. Die gesicherte empirische Wirkung überträgt sozusagen die positive Wertung der Oberprämisse prosoziale Einstellung auf die auslösende Bedingung der gesellschaftlichen Anerkennung des spezifischen Opferschicksals. Diese Anerkennung ist damit unter Bezug auf die übergeordnete Positiv-Wertung von Prosozialität (relativ) legitimiert. Durch das Aufsteigen zu immer generelleren präskriptiven Oberprämissen werden schnell die Meso- und Makroebenen erreicht (in diesem Fall z. B. allgemeine Menschenrechte als Grundlage für prosoziale Einstellungen gegenüber allen Menschen unterschiedlichster Ethnizität, Religionszugehörigkeit etc.), die für Pädagogik, Soziologie, Politologie etc. konstitutiv sind. Eine rationale Diskussion und Begründung der präskriptiven wissenschaftlichen Thesen einer Friedenspsychologie ist daher letztlich nur in interdisziplinärer Kooperation mit diesen anderen sozialwissenschaftlichen Disziplinen möglich. Durch eine solche systematische Wertungsanalyse lassen sich vor allem implizite Wertungsgewohnheiten überwinden, die sich wegen des WFP und des daraus folgenden Verzichts auf rationale Wertdiskussionen unbemerkt in die epistemische Forschung einschleichen (können).

So haben sich Alltagsvorstellungen wie "Gutes kommt (nur) aus Gutem" bzw. "Schlechtes führt (immer) zu Schlechtem" auch in sozialpsychologischen Modellen zu Engels- bzw. Teufelskreisen festgesetzt, etwa dem Teufelskreis der Gewaltspiralen, während die friedenspsychologische Modellierung wertungsreflektierend konstruktiv nach den Möglichkeiten fragt, wie aus negativen Bedingungen (Opfer-

erlebnisse und -bewusstsein) positive Wirkungen (von Solidarität und Versöhnungsbereitschaft) erwachsen können (z. B. Vollhardt 2009b).

Nicht zuletzt legt die systematische interdisziplinäre Wertungsanalyse aber auch die Grundlage für Modelle, wie die elaborierten Ziele (hier des Friedens) in die Alltagsrealität zu implementieren sind. Das umfasst eine entsprechende Implementierungsforschung – in der friedensfördernde Maßnahmen, die auf psychologischen Theorien beruhen, nicht nur in Interventionsprogramme umgesetzt werden, sondern in der die Wirksamkeit dieser Programme auch überprüft wird (vgl. z. B. Paluck 2009) – als letzter, komplettierender Schritt einer umfassenden praktischen Friedensforschung!

#### Elemente einer konstruktivistischen Konflikttheorie

Heinz Messmer

"Eine Wirklichkeit, welche die Subjektivität transzendiert, indem sie in solchen Aussagen dargestellt wird, deren Wahrheit anhand von explizit vereinbarten und methodisch begründeten Regeln (= transsubjektiv) nachgeprüft werden kann, ... können wir auch als *objektive Wirklichkeit* bezeichnen."

Kempf 2009b, S. 61

#### 1 Zur Konzeptualisierung einer Theorie sozialer Konflikte

Worüber sprechen wir, wenn von sozialen Konflikten die Rede ist? Welcher Aspekt sozialer Wirklichkeit ist damit bezeichnet? Was zeichnet den Konflikt substanziell aus? Wenn man in den herkömmlichen Konflikttheorien nach Antworten auf diese und ähnliche Fragen sucht, stößt man schon bald auf unterschiedliche Perspektiven: Je nachdem, aus welchem Blickwinkel die jeweilige Theorie ihren Gegenstand in Augenschein nimmt, kommen verschiedene Aspekte einer Konfliktwirklichkeit in Betracht. Einen working consensus der konflikttheoretischen Forschung dagegen, ihre analytische Einheit mithin, findet man nur auf einer axiomatischen Basis begründet: Was den Konflikt von allen anderen Tatsachen des Sozialen unterscheidet, ist die Modalität der Unvereinbarkeit, der Negation; ist Gegensatz bzw. Ausschließlichkeit im Verhältnis wenigstens zweier Standpunkte oder Perspektiven (vgl. Deutsch 1990). Jenseits dieser Basisdefinition werden je nach theoretischer Strömung und Zielsetzung mitunter sehr verschiedene Ausgangspunkte markiert, die unterschiedliche Seiten der Konfliktwirklichkeit zu Tage fördern. Das hat nicht nur Auswirkungen für die Art einer Konfliktkonzeption, sondern macht den Betrachter auch für Einseitigkeiten der Theorie sensibel. Wenn beispielsweise spieltheoretische Modelle primär die Ratio des (kommunikationslosen) Entscheidungskonflikts repräsentieren (vgl. selbstkritisch Rapoport 1974, S. 224f.) oder das Konzept des character contest die Anerkennungsaspekte des Beziehungskonflikts reflektiert (so etwa Goffman 1959), ist damit noch nicht gesagt, dass die Angemessenheit der einen Perspektive zugleich auch den anderen Konflikttypus zutreffend bezeichnet.

Die hier entwickelte Sichtweise reagiert auf verschiedene Inkonsistenzen der konflikttheoretischen Forschung. Der Hauptkritikpunkt besteht dabei darin, dass herkömmliche Konflikttheorien ihre Haltepunkte vorwiegend *außerhalb* ihres Gegen-

standsbereichs suchen. Was einen Konflikt im Rahmen eines konventionellen Paradigmas substanziell auszeichnet und inhaltlich charakterisiert, sind vornehmlich seine Auslöseanlässe: so zum Beispiel individuelle Bedürfnisse und Interessen, Nutzen- und Machterwägungen von Gruppen und Organisationen oder eine nationalstaatliche bzw. globale Ungleichverteilung von Lebenschancen u. ä. m. Sofern die konventionelle Konfliktforschung hauptsächlich auf konflikt*induzierende* Einflüsse fokussiert, bleibt der Konflikt als solcher dahinter verborgen.

Nach Dahrendorf (1971, S. 109) beruht der soziale Konflikt beispielsweise auf an Herrschaft orientierten gesellschaftlichen Strukturarrangements und geht zwingend daraus hervor. Alle Konfliktbildungen, die diesem Strukturbegriff nicht zuordenbar sind, fallen diesem Raster zum Opfer. Die Gleichsetzung von Konflikt mit den Kontexten seiner Entstehung begünstigt weit reichende Einseitigkeiten und formal zweifelhafte Postulate der Theorie. Empirisch lässt sich zwar zeigen, dass ein Konflikt bestimmte Ursachen in bestimmte Wirkungen transformiert, jedoch wird keineswegs deutlich, wie sich der Prozess der Umwandlung praktisch vollzieht. Entsprechend kommt der Konflikt allenfalls als ein statisches Phänomen in den Blick, als eine Art trivial konzipierter Blackbox, die gleichsam mechanisch die Impulse aus ihrer Umwelt in einen Gegensatz transformiert.

Eine von konkreten Kontexten unabhängige und diese damit übergreifende Konflikttheorie steht deshalb vor dem Problem, einer analytischen Orientierung zu genügen, die einerseits nicht zu konkret optiert (um das Einzelne nicht mit dem Allgemeinen zu verwechseln), die andererseits aber auch dazu imstande ist, den empirischen Konfliktsachverhalten Rechnung zu tragen. An die Stelle herkömmlicher Kontext- bzw. Kausalannahmen tritt die Idee einer systemischen Form, die zwischen dem Konflikt und den Kontexten seiner Entstehung strikt unterscheidet. Damit wird die in der Konflikttheorie gängige Frage nach den (externen) Ursachen seiner Entstehung ersetzt durch die Frage, auf Grund welcher internen Prozesse und Strukturen der Konflikt seine Formen gestaltet und variiert.

Vor diesem Hintergrund wird nachfolgend die These vertreten, dass mit Hilfe einer konstruktivistischen Theorieperspektive die Kluft zwischen internen und externen Bedingungen der Konfliktkonstitution nicht nur theoretisch, sondern auch empirisch überbrückt werden kann. Im weitesten Sinne bezeichnet der Begriff Konstruktivismus ein Beschreibungs- und Erklärungsmodell, demzufolge die Sachverhalte des Sozialen ausnahmslos das Ergebnis zwischenmenschlicher Interaktionen sind. Empirisch liegt der Prüfstein sozialer Wirklichkeitskonstruktionen in den Kommunikationen der Gesellschaftsmitglieder begründet. Berger und Luckmann beispielsweise sprechen in diesem Zusammenhang von der "wirklichkeitsstiftenden Macht des Gesprächs" (1969, S. 164) bzw. vom "Rattern einer Konversationsmaschine" (ebd., S. 163), mittels derer der Einzelne unentwegt auf seine subjektiven Wirklichkeitsannahmen rekurriert und diese im Vollzug der Interaktion mit anderen verfestigt oder modifiziert.

Entsprechendes gilt für den Konflikt. Analog zur Ausdifferenzierung sozialer Systeme (vgl. Luhmann 1984) lassen sich dessen Form- und Struktureigenschaften ent-

30 Heinz Messmer

lang den Selektionsbedingungen der Kommunikation als den nicht mehr hintergehbaren Letzt-Elementen sozialer Interaktionen beschreiben. Alles, was als Konfliktwirklichkeit den Status einer empirischen Tatsache zugesprochen bekommt, muss demgemäß durch das Nadelöhr wechselseitig aufeinander Bezug nehmender Kommunikation hindurchgeschleust werden. Denn nur so ist sichergestellt, dass ein anderer die Unvereinbarkeit wahrnehmen und darauf Bezug nehmen kann. Der Konflikt wird somit als ein kommunikatives Geschehen betrachtet, das auf der Basis von Kommunikationen entsteht, prozessiert und letzten Endes auch wieder verschwindet.

Die Einführung einer konstruktivistischen Perspektive in die Konfliktanalyse ist in gewisser Weise ein redundantes Unterfangen, da der soziale Konflikt ohnehin nur beobachterabhängig existiert, weil er von den Deutungen und Deutungssystemen der Beteiligten abhängig ist. Darauf komme ich noch näher zu sprechen. Die Etablierung einer konstruktivistischen Perspektive lenkt jedoch die Aufmerksamkeit explizit auf den Sachverhalt, dass vor allem solche Einflüsse konfliktwirksam sind, die aus den (beobachtungsabhängigen) Interaktionen der Konfliktbeteiligten selbst resultieren. Deshalb muss jede Forschung, die sich eine konstruktivistische Sichtweise zu eigen macht, ihre Analysen auf die Verkettung empirischer Interaktionen konzentrieren.

Im Folgenden möchte ich einige Elemente skizzieren, die aus meiner Sicht notwendig sind, um eine konstruktivistische Konfliktperspektive empirisch robust zu verankern. Dabei bediene ich mich vor allem basaler Theoriebausteine der funktionalistischen Systemtheorie (vgl. Luhmann 1984), die sich über beobachtungsabhängige Phänomene begründen (Kap. 2). Auf dieser Grundlage werde ich die basalen Prozesse der Ausdifferenzierung einer Konfliktstruktur näher erläutern (Kap. 3) und Retrospektion und Interpunktion als maßgebliche Einflussfaktoren auf die Stabilisierung einer für den Konflikt unverzichtbaren Perspektivendifferenz diskutieren (Kap. 4). Mit einigen weiterführenden Ausblicken werde ich diese Darstellung schließen (Kap. 5).

#### 2 Elemente einer konstruktivistischen Konflikttheorie

Der Ausgangspunkt konstruktivistischer Analysen besteht in der Überlegung, dass die (Wirklichkeits-)Konstruktion sozialer Konflikte auf der beobachtungsabhängigen Verkettung von Interaktionen beruht. Ohne Interaktion, an der sich ein wie auch immer gearteter Widerspruch manifestiert, ist der soziale Konflikt praktisch undenkbar. Interaktionen erweisen sich mithin als der analytische Kern einer konstruktivistischen Perspektive auf den sozialen Konflikt. Mit den nachfolgenden Überlegungen werden diejenigen Eigenschaften sozialer Interaktion thematisiert, die sich für eine konstruktivistische Orientierung in der Konfliktforschung als besonders bedeutsam erweisen.

#### 2.1 Doppelte Kontingenz

Eine Interaktion ist der Versuch zweier (oder mehrerer) Individuen, ihr Verhalten zu koordinieren und sinnhaft aufeinander zu beziehen, sobald sie zueinander in Kontakt treten. Das daraus resultierende Problem wird seit Parsons (1968) unter dem Begriff der doppelten Kontingenz diskutiert. Unter den Begriff Kontingenz fallen alle diejenigen sozialen Sachverhalte, die weder notwendig noch unmöglich sind. Dass beispielsweise auf den Gruß ein Gegengruß folgt, ist normativ zwar wahrscheinlich, empirisch aber keineswegs sicher. In der Interaktion tritt Kontingenz immer doppelt bzw. beidseitig auf, auf Seiten von Alter ebenso wie auf Seiten von Ego. Beide sind in der Wahl ihrer Verhaltensalternativen grundsätzlich frei in dem Sinne, dass sie selbst darüber entscheiden, welche Einflüsse für die Wahl ihres Verhaltens maßgeblich sind.

Jede Interaktion löst (und stellt!) das Problem doppelter Kontingenz von Augenblick zu Augenblick neu. Eine Verhaltensfestlegung hätte auch anders ausfallen können, jedoch haben sich die Beteiligten für dieses (und nicht für ein anderes) Verhalten entschieden. Folgerichtig hebt diese Begrifflichkeit die Selektivität allen Verhaltens hervor. Sie stellt klar, dass jede Interaktion sich als das (kontingente) Produkt einer Wahl aus unterschiedlichen Verhaltensalternativen präsentiert, die auch anders hätte getroffen werden können. Unter dem Blickpunkt doppelter Kontingenz tritt Interaktion deshalb als ein kontinuierlich zu lösendes Problem in Erscheinung: Wie gelingt es Alter und Ego, ihr Verhalten im Bewusstsein ihrer wechselseitigen Abhängigkeit auszuwählen und zu koordinieren? Welche Voraussetzungen sind dafür maßgeblich?

#### 2.2 Wahrnehmung/Beobachtung

Eine erste Voraussetzung für die Lösung des Problems liegt in der Wahrnehmungsabhängigkeit sozialen Verhaltens. Interaktion kommt nur zustande, wenn beide Seiten auf eine (wie auch immer geartete) Form der wechselseitigen Wahrnehmung zurückgreifen können, um ihr Verhalten zu koordinieren. Ohne Wahrnehmung keine Interaktion. Wahrnehmung vermittelt Informationen über den anderen, die mitentscheidend für die eigenen Orientierungen und situativen Verhaltensfestlegungen sind.

Den Gesichtspunkt der Wahrnehmung (oder Kognition) hat die luhmannsche Systemtheorie im Anschluss an die biologistische Epistemologie von Autopoiesis als Beobachtung reformuliert (vgl. Luhmann 1995, S. 92ff.). Beobachten ist definiert als ein Prozess des Bezeichnens im Kontext des Unterscheidens. Man sieht beispielsweise den Entgegenkommenden auf der Straße, erkennt ihn als Fremden und tritt zur Seite, damit er – ohne sich bedroht zu fühlen – vorbeigehen kann. Dasselbe Verhalten einem Bekannten gegenüber dagegen befremdet. Unterscheidungen nach dem Schema bekannt/unbekannt verändern die Definition einer Situation hinsichtlich ihrer Konsequenzen in bedeutsamer Weise. Bezogen auf den Konflikt und die Prozesse der Konflikteskalation üben die Wahrnehmungen der Beteiligten

32 Heinz Messmer

einen maßgeblichen Einfluss aus; vor allem dann, wenn sie sich an den Konfliktverlauf adaptieren. Folgerichtig definiert Kriesberg (1973, S. 17) den sozialen Konflikt auf der Basis einer *subjektiv* wahrgenommenen Unvereinbarkeit und betrachtet das subjektive Erleben der Beteiligten (awareness) als eine maßgebliche Größe der Konfliktkonstitution.

#### 2.3 Sinn

Gemäß diesen Überlegungen ist die bloße Wahrnehmung (bzw. Beobachtung) weder ein Automatismus noch ein Garant für koordiniertes Verhalten. Weder geschieht sie mechanisch noch neutral oder werturteilsfrei. Vielmehr unterliegt sie einem sinnhaften Bearbeitungsprozess, der die Sachverhalte der Wahrnehmung in das individuelle Beobachtungsraster integriert. Wahrnehmung *ist* Konstruktion, weil der Wahrnehmungsgegenstand seine Bedeutung nicht schon in sich trägt, sondern sinnhaft bezeichnet und bewusst werden muss (vgl. Luhmann 1984). Mittels bestehender Bewusstseinsstrukturen (Wissen, Erfahrung, Gedächtnis) wird Wahrnehmung in ein sinnhaftes Erleben transformiert, das alternative Verweismöglichkeiten eingrenzt und spezifische Verhaltensoptionen in den Vordergrund rückt. Für den einen sind beispielsweise bestimmte Merkmale im Aussehen oder Verhalten des anderen Grund dafür, ihn zu meiden; für den zweiten hingegen Ansporn für Zuwendung und Kontakt.

Obwohl in einer jeden Wahrnehmung ein gewisses Maß an Übereinstimmung (oder Objektivitätsbezug) vorausgesetzt ist, sind gleichläufige Subjekt-Objekt-Unterscheidungen bei der Interaktion keineswegs sicher. Wenn das Objekt der Wahrnehmung Gegenstand sinnhaften Erlebens wird, lässt sich zwischen Objekt und Subjekt bzw. Beobachtetem und Beobachter nicht mehr hinreichend genau unterscheiden. Aus diesen Gründen hatte schon die klassische Phänomenologie kausale Begründungslogiken vom Kopf auf die Füße gestellt, indem sie nicht danach fragt, was Wirklichkeit objektiv *ist*, sondern unter welchen Bedingungen Individuen ihre Beobachtungen für objektiv *wirklich halten* (so bereits James 1890).

Wahrnehmung ist insofern eine notwendige, aber keine hinreichende Bedingung gelingender Interaktion. Vielmehr ist sie ein Produkt der selektiven Aufmerksamkeit, des subjektiven Interesses und nicht zuletzt auch der Verträglichkeit der Wahrnehmung mit bereits verfügbarem Wissen. Sie kann vom Prozess der Sinnbearbeitung nicht abgelöst werden. Umso mehr stellt sich die Frage, wie das Problem doppelter Kontingenz gelöst werden kann, zumal mit den Prozessen der Interaktion auch die Strukturen der Sinnselektion kontinuierlich variieren.

#### 2.4 Struktur (Positivversion doppelter Kontingenz)

Situationen doppelter Kontingenz setzen ein Mindestmaß an wechselseitiger Beobachtung voraus, ohne vollständige Transparenz (im Sinne von Durchschaubarkeit der Motive) zu erreichen. Gerade im Hinblick auf den Anfang einer noch (verhältnismäßig) voraussetzungslosen und unstrukturierten Interaktionssituation stellt sich dieses Problem somit in voller Schärfe. Im Kontakt zum anderen ergibt sich anfänglich die Notwendigkeit, dass die Beteiligten ihre Absichten füreinander produktiv und greifbar gestalten.

In vergleichbarer Weise hat Niklas Luhmann (1984, S. 166ff.) den Strukturaufbau in doppelt kontingenten Situationen zirkulär formuliert: Ich tue, was Du willst, wenn Du tust, was ich will. Er spricht von der Positivversion doppelter Kontingenz als einer mehr oder weniger verlässlichen Kernstruktur sozialer Systeme. Aber auch dies bezeichnet allenfalls eine Erwartungsstruktur, die als solche erst einmal kenntlich gemacht und interaktiv validiert werden muss. Symbolische Formen der Kommunikation (z. B. ein Lächeln) und/oder sprachliche Explikationen ("Darf ich mich vorstellen?") machen die Absichten der Beteiligten zumindest ansatzweise durchschaubar und vermitteln dem anderen einen vorläufigen Anhaltspunkt, wie das Problem eines ersten Zuges enttautologisiert und überbrückt werden kann.

Um unter Bedingungen prinzipieller Unsicherheit erste Züge zu initiieren, haben Gesellschaften normative Erwartungsstrukturen etabliert, die dafür sorgen, dass eine Interaktion in die gewünschte Richtung verläuft. Dazu gehören beispielsweise Takt, Höflichkeit und Vertrauen. Empirisch lässt sich unschwer belegen, dass sich die Sozialwelt vorzugsweise auf Erwartungsstrukturen stützt, welche die Positivversion doppelter Kontingenz normalisieren (vgl. näher Brown & Levinson 1987). Im Grundsatz ist unser Sozialverhalten an kooperativen Verhaltensvorstellungen orientiert, während davon abweichende Verhaltenserwartungen einer besonderen Begründung bedürfen. Dabei spielt nicht zuletzt die Abwägung von Kurz- und Langfristinteressen eine wichtige Rolle. Jede Interaktion stellt die Beteiligten vor die Notwendigkeit, die eigenen Absichten und Interessen mit denen des anderen so zu verknüpfen, dass eine wechselseitige Zusammenarbeit möglich wird. Wie die spieltheoretischen Studien von Axelrod (1988) zeigen, sind insbesondere solche Verhaltensorientierungen erfolgreich, die auf Kooperation und Wechselseitigkeit basieren, im Falle der Defektion jedoch Gleiches mit Gleichem vergelten (TIT FOR TAT). Langfristig zahlt sich das anfängliche Risiko (Kooperation im ersten Zug) anscheinend aus, selbst wenn die Motive und Verhaltensabsichten von anderen bis dahin noch diffus und unklar sind.

#### 2.5 Prozess

Zweifellos existieren Takt, Höflichkeit oder Vertrauen als Ausdruck einer gesellschaftlich relevanten Erwartungsstruktur. Inwieweit sie tatsächlich zur Anwendung kommen, ist jedoch weniger eine normative als eine empirische Frage. Entsprechend lässt sich der Strukturaufbau interaktiver Orientierungen nur dann angemessen begreifen, wenn man die ihm zugrunde liegenden Prozesse abschätzen kann.

Was aber ist ein Prozess? Wer in den Sozialwissenschaften nach eigenständigen und generalisierbaren Prozesstheorien sucht, wird kaum eine finden. Sozialwissenschaftlich prozessorientiertes Denken geschieht vorwiegend im Fahrwasser der 34 Heinz Messmer

darwinschen Evolutionstheorie, zuweilen auch im Kontext kybernetischer Kreislaufmodelle. Unter Umständen werden sozialwissenschaftliche Erklärungsmodelle unter Prozessgesichtspunkten reformuliert (vgl. Miebach 2009). Mit Blick auf Interaktion ist das Reiz-Reaktions-Modell ein in die Jahre gekommenes, aber für die Erklärung sozialer Konflikte immer noch attraktives Beschreibungsmodell. Es geht von der grundsätzlichen Annahme aus, dass Interaktion Reiz und Reaktion gleichzeitig ist: Indem Ego auf Alters Verhalten reagiert, ist seine Reaktion gleichzeitig Anreiz für Alters Verhalten (vgl. Bateson & Jackson 1964). Dieses Modell leuchtet ein, da Interaktion sequenziell bzw. zugtypisch geordnet abläuft. Für eine konstruktivistische Konflikttheorie ist es relevant, weil es die Voraussetzungen aufzeigen kann, unter denen die Beteiligten im Konflikt zirkuläre Verhaltensketten zum Zwecke gegenläufiger Ursachenzuschreibungen verfestigen.

Rein logisch betrachtet sind Prozesse der Interaktion dem Strukturaufbau vorgelagert, faktisch jedoch verhalten sich Struktur und Prozess zirkulär. Prozesse nehmen in Anspruch, was sie selbst produzieren; so zum Beispiel, wenn sich die Beteiligten an unterstellten Erwartungen orientieren, diese durch ihr Verhalten zum Ausdruck bringen und dadurch reifizieren. Am Prozess orientierte Analysen sind für eine konstruktivistische Perspektive aufschlussreich, weil sie die Ausdifferenzierung sozialer Wirklichkeitskonstruktionen zu rekonstruieren erlauben. Auf dieser Grundlage ergeben sich für die Beobachtung sozialer Konflikte mithin innovative und weiterführende Fragen: Wie können Konflikte aus kooperativen Ordnungen emergieren? An welche Voraussetzungen sind die Prozesse der Ausdifferenzierung sozialer Konflikte geknüpft? Welches sind ihre unverzichtbaren Ingredienzien? Und inwieweit verlaufen solche Entwicklungen konsistent und linear?

#### 2.6 Kommunikation

Sehr viel klarer zeigt sich dieser Zusammenhang, wenn man dem Umstand Rechnung trägt, dass sich die Interaktionen von Alter und Ego auf Kommunikationen gründen. Wie bereits angeführt, lassen sich ab dem Moment, in dem Alter und Ego wechselseitig in ihr Wahrnehmungsfeld treten, aus ihrem Verhalten weit reichende Rückschlüsse ziehen. Deshalb lautet eines der bekanntesten Axiome der pragmatischen Kommunikationstheorie, dass man nicht nicht kommunizieren kann (vgl. Watzlawick et al. 1982).

Luhmann (1984) definiert Kommunikation als das Zusammenwirken von Mitteilung, Information und Verstehen. Diese Definition macht auf eine Reihe weit reichender Implikationen aufmerksam. Anders als Handlung oder Verhalten ist eine Kommunikation als Ganzes nicht individuell zurechenbar, allenfalls nur ein Teil von ihr, die Mitteilung. Sie ist insofern auf die Mitwirkung beider Seiten angewiesen, eine, die sich mitteilt und eine, die versteht – wodurch nahegelegt wird, die Begriffe Interaktion und Kommunikation synonym zu gebrauchen. Derselbe Umstand rechtfertigt es im Übrigen auch, bereits auf der Ebene einer einzelnen Kommunikation von einem sequenziellen Ablauf zu sprechen: Unabhängig von ihrer Kürze und Dauer geht dem Verstehen die Mitteilung der Information voraus, was zwei

logisch aufeinander folgende Schritte impliziert. Kommunikation wird – drittens – schließlich zu einem Prozess, wenn derjenige, der versteht, an sein Verstehen mit einer eigenen Mitteilung anschließt (was vor dem Hintergrund des Axioms der pragmatischen Kommunikationstheorie praktisch nicht zu vermeiden ist).

Die Konsequenzen der analytischen Umstellung von Verhalten auf Kommunikation sind vielfältig und komplex und für eine konstruktivistische Sichtweise ungemein inspirierend. Denn gegenüber dem Handlungsbegriff weckt der der Kommunikation deutlich mehr dynamische Assoziationen. Kommunikation verknüpft Anfang (Mitteilung) und Ende (Verstehen) ebenso wie ihre Protagonisten (Ego und Alter) zu einem fließenden Ablauf, in dem Anschlussschritte unausweichlich erscheinen, so dass nicht zuletzt auch das Ende einer Kommunikation erklärungsbedürftig wird. Denn anders als Handlung verfügt Kommunikation weder über ein natürliches Ende noch über ein Ziel, an dem sich ihr Ende ablesen ließe.

#### 3 Konfliktemergenz

Nach dem bisher Gesagten wird deutlich, dass Interaktion trotz ihrer Zwangsläufigkeit durchaus nicht selbstverständlich ist, ihre Realisierung vielmehr komplexen Voraussetzungen unterliegt, die erfüllt sein müssen, damit sie gelingen und fortschreiten kann. Vor diesem Hintergrund liegt schließlich die Vorstellung nahe, dass sich der soziale Konflikt als Abweichung von auf Kooperation abgestellten Interaktionsbeziehungen ausdifferenziert: Demnach ist er das Resultat des Zusammenwirkens wenigstens zweier Individuen, für die die Positivversion doppelter Kontingenz nicht mehr maßgebend ist. Die Generierung sozialer Konflikte bleibt jedoch auf entsprechende Äußerungsformen angewiesen, damit ein Gegensatz wahrnehmbar und interaktiv wirksam werden kann.

In diesem Zusammenhang fungieren "Ja" und "Nein" als elementare Bestandteile der Kommunikation, die über Annahme oder Ablehnung mitgeteilter Informationen entscheiden. Jeder Interaktion ist mithin die Wahl zwischen Kooperation oder Konflikt inhärent. Ebenso wie Kooperation werden Konflikte im Rahmen von Kommunikation generiert, transformiert und wieder beendet. Es reicht daher nicht aus, die Existenz des Konflikts (wie in den Sozialwissenschaften weithin üblich) über unterstellte Interessengegensätze zu postulieren. Aus einem Interessengegensatz resultiert nicht notwendig Konflikt – wenigstens nicht, solange sich die Interaktionen der Beteiligten nicht auch füreinander erkennbar an einem Widerspruch orientieren.

Die Realität des Konflikts benötigt als Bedingung seiner Möglichkeit das empirische Nein oder äquivalente Ausdrucksformen, damit der Widerspruch für die Beteiligten greifbar wird. Diese Voraussetzung allein begründet aber noch nicht die Existenz des Konflikts. Denn auch wenn Ego der Mitteilung Alters widerspricht, bleibt bis auf Weiteres offen, ob und inwieweit ein Alter Egos Widerspruch akzeptiert oder seinerseits ablehnt. Grundlage für die Umpolung einer prinzipiell kooperativen Interaktionsbeziehung (Positivversion doppelter Kontingenz) ist vielmehr die wech-

36 Heinz Messmer

selseitige Kenntlichmachung von Unnachgiebigkeit. Erst wenn Alter trotz Widerspruch an seiner (ursprünglichen) Sinnzumutung festhält und Ego trotz Gegenwiderspruch auf Ablehnung insistiert, tendiert der Konflikt zu einer interaktiv stabilen Struktur, die sich als Negativversion doppelter Kontingenz ausdifferenziert und dafür sorgt, dass die Interaktion unter negativen Vorzeichen fortgesetzt wird: Ich tue nicht, was du möchtest, wenn du nicht tust, was ich möchte (vgl. Luhmann 1984, S. 531). Wenn daher die Unnachgiebigkeit des einen auf die des anderen trifft, kommt Kooperation rasch zum Erliegen. An ihre Stelle tritt eine Orientierung, welche die wechselseitige Ablehnung von Sinnzumutungen in den Brennpunkt der Aufmerksamkeit rückt. Diese Entwicklung verdeutlicht Abbildung 1.

	Widerspruch	Konflikt	Opposition	Struktur
1. Zug (Alter)	Sinn			
2. Zug (Ego)	Gegensinn <u></u>	> Widerspruch		
3. Zug (Alter)		widerspruch ⊏	⇒ Unnachgiebigkeit	
4. Zug (Ego)			Unnachgiebigkeit	Negativ- Version doppelter Kontingenz

Abbildung 1: Struktur und Prozess. Ausdifferenzierung eines Konfliktsystems (vgl. Messmer 2003)

Nach dieser Darstellung handelt es sich beim sozialen Konflikt um eine Aneinanderreihung widersprüchlicher Synthesen einer Sinnselektion. Eine Sinnzumutung ist demnach nur dann konfliktrelevant, wenn sie auf Gegensinn trifft, während ein Widerspruch nur prozessieren kann, sofern er seinerseits auf Widerspruch stößt. Jeder einzelne Zug in dieser Kette ist auf die Wahrnehmung einer Unvereinbarkeit angewiesen, ohne die sich der Konflikt nicht vollständig ausdifferenzieren kann. Die Ausdifferenzierung des Konflikts ist also im Kern ein Produkt seiner internen Differenzierung mit den Mitteln der Widerspruchskommunikation, die weitere Konstruktionsleistungen nach sich zieht.

#### 4 Perspektivendivergenz

Eine in der Konflikttheorie weniger häufig elaborierte Frage zielt auf das Zustandekommen und Prozessieren von Perspektiveninkongruenz. Typischerweise ist unterstellt, dass Individuen infolge unterschiedlicher Interessen und Werte gegenläufige Ansichten und Positionen verfolgen, was die Fortführung des Konflikts

jedoch nur zum Teil erklärt. Gemessen an den meist hohen sozialen Kosten einer Konflikteskalation liegt die Vermutung nahe, dass Nachgeben oder Kompromiss bevorzugte Varianten der Konfliktbeilegung sind (für empirische Nachweise vgl. beispielsweise Vuchinich 1987, Keppler 1994). Daraus resultiert die Frage, wie der Konflikt die Bedingungen seiner eigenen Fortsetzbarkeit generiert.

# 4.1 Retrospektion

Ein erster Baustein betont die retrospektive Verdopplung einer ko-produzierten Sinnkonstruktion. Wenn beispielsweise Alters Äußerung auf Egos Widerspruch trifft, wird die positive Wertigkeit einer Sinnzumutung durch eine negative Wertigkeit dupliziert. Der Widerspruch verdoppelt mithin eine Sinnperspektive unter dem Gesichtspunkt ihrer Ablehnungswürdigkeit. Auf diese Weise entsteht ein Verhältnis logischer Inkongruenz ( $A \neq A$ ), was notwendigerweise Problemlösungsaktivitäten impliziert, sofern klargestellt gestellt werden muss, welche Orientierung (A oder  $\neq A$ ) im Weiteren Gültigkeit hat (vgl. Luhmann 1984, S. 493 ff.).

Obwohl also die Ablehnung einer Sinnzumutung erstmalig Konflikt indiziert, ist sie mit dem Konfliktanfang nicht identisch. Vielmehr wird der Auslöseanlass des Konflikts (Sinnzumutung als erster Zug) erst durch den Widerspruch als dem zweiten Zug in der Kommunikation *retrospektiv* sichtbar. Was immer der Widerspruch sonst noch beabsichtigt: Es geschieht mit Blick auf eine ihm voraus liegende Aktivität. Sequenzanalytisch betrachtet ist die erste Zugposition im Konflikt formal das Ergebnis der zweiten, die insofern ihren eigenen Anlass mitkonstruiert, ihn aber gleichzeitig nicht selbst verantworten muss.

Die Folgen dieser zeitlichen Asynchronizität für den Konfliktverlauf können gar nicht hoch genug eingeschätzt werden. Denn damit wird eine Situation virulent, in der die Konfliktbeteiligten hinsichtlich der Ursachen und Auslöseanlässe des Konflikts unterschiedliche Standpunkte einnehmen können: Aus ihrer Sicht ist es immer der andere, der mit seinen Aktivitäten den Konflikt perpetuiert.

# 4.2 Interpunktion

Ein mit diesen Überlegungen eng verknüpftes Konzept ist das der Interpunktion, das auf die Folgen retrospektiver Dopplungen aufmerksam macht. Interpunktion steht für den Versuch, zirkulär verkettete Ereignisabfolgen mittels eines Kausalschemas zu zerlegen (vgl. Watzlawick et al. 1982, S. 57ff.). Damit wird vor allem die Stabilität regenerativer Konfliktzirkel evident: Denn obwohl ein jedes Verhalten Reiz und Reaktion gleichzeitig ist, werden die Verhaltensketten im Konflikt von den Beteiligten in ungleich strukturierte Triaden unterteilt, so dass widersprüchliche Einschätzungen darüber entstehen, was Ursache und Wirkung einer Verhaltenssequenz ist.

In Beziehungskonflikten werden Verhaltensketten typischerweise so interpunktiert, dass das Verhalten des jeweils anderen am Anfang einer Triade (Reiz/Reak-

38 Heinz Messmer

tion/Verstärkung) steht, also ursächlich für das eigene Handeln ist. Dieses erscheint damit als Reaktion auf das Verhalten des anderen, was es erlaubt, die Verantwortung für das Konfliktgeschehen auf die gegnerische Seite zu projizieren. Darüber hinaus wird das eigene Verhalten als logisch konsistent, das des anderen hingegen als damit unvereinbar erlebt. Durch Interpunktion wird die Negativversion doppelter Kontingenz für beide Seiten verbindlich, sofern immer der andere durch sein Verhalten die eigene Unnachgiebigkeit strukturell reproduziert. Wilhelm Kempf (1993) spricht in diesem Zusammenhang von einer permanenten Reproduktion der konfliktauslösenden Konstellationen. Dabei resultiere die objektive Bedeutung von Konfliktlösungsversuchen für das System aus der "eigentümlichen Logik des Konflikts selbst" (S. 61).

Einander widersprechende Kausalzurechnungen in der Konfliktkommunikation sind in dem Maße folgenreich, wie damit die Perspektivendivergenz im Konflikt – und damit einhergehend die Positionsdifferenzen – selbstrekursiv kontinuieren. Kommunikation im Konflikt wird mithin zum Produkt einer epistemischen Suche nach den Ursachen einer negativen Erfahrung, die entsprechende Erklärungsversuche nach sich zieht: Was ist die Ursache? Wer ist verantwortlich? Wen trifft die Schuld?

Perspektiveninkongruenz kommt nach diesen Überlegungen dadurch zustande, dass ein Konfliktsachverhalt zwar ko-produziert und beidseitig ist, diesbezügliche Verantwortlichkeitszuweisungen jedoch einseitig und gegenläufig prozessieren. Auf dieser Grundlage wird das Konflikterleben in verschiedenen Sinndimensionen kommuniziert: In der Sachdimension orientiert sich der Konflikt am Dual von Schaden und Nutzen, in der Sozialdimension am Dual von Ich und Du, in der Zeitdimension schließlich am Dual von Anfang und Ende. Damit werden hinreichend vielfältige Anknüpfungspunkte generiert, damit ein Konflikt diejenigen Sachverhalte und Themen hervorbringen kann, die ihm eine rekursive Weiterentwicklung auf Dauer ermöglichen.

# 5 Konsequenzen für eine konstruktivistische Konfliktforschung

In der hier vertretenen Sichtweise sind Konflikte primär das Produkt einer sequenziellen Abfolge der Widerspruchskommunikation. In dem Maße, wie der Konflikt prozessiert und sich kommunikativ ausdifferenziert, ordnet er Personen und Themen nach einer Erwartungsstruktur, welche vorzugsweise die Fortführung des Konflikts unterstützt. Aus der Binnensicht könnte man so formulieren: Konflikte erzeugen selbstreferenziell diejenigen Individuen und Umwelten, die zu ihrer Fortführung notwendig sind, während alle Aspekte, die nicht dazu passen, mit zunehmender Konfliktentwicklung an Bedeutung verlieren.

Für die Frage, welche Bedeutung den Beobachterperspektiven der Beteiligten im Konflikt eingeräumt werden muss, lassen sich die hier skizzierten Konflikteigenschaften mithin zu einer allgemeinen Regel verdichten: Mit fortschreitender Konfliktentwicklung nimmt die Wahrscheinlichkeit zu, dass die Beteiligten sich aus-

schließlich auf den 'Eigensinn' ihrer Wahrnehmungsmuster verlassen, die für Fremdverstehen oder Korrekturen immer weniger zugänglich sind. Im Zuge einer Konflikteskalation orientiert sich das 'Verstehen' zunehmend weniger an der wechselseitigen Ergründung von Standpunkten und Perspektiven, sondern stattdessen an der kognitiven Absicherung und Immunisierung der eigenen Wahrnehmungswelt. Morton Deutsch (1973, S. 365) hat diesen Konnex in *Deutsch's crude law of social relations* treffend zusammengefasst: "the characteristic processes and effects elicited by an given type of social relationship also tend to elicit that type of social relationship".

Diesen Einsichten zufolge basiert eine konstruktivistische Konfliktperspektive paradoxerweise zuerst und vor allem auf Dekonstruktion: Sie hinterfragt die selbstverständlichen Annahmen und Prämissen herkömmlicher Perspektiven, verzichtet auf epistemologisch vorschnelle Reduktionen und zerlegt wissenschaftlich eingeschliffene Begriffe und Kategorien in empirisch begründete Dimensionen der sozialen Sinnproduktion. Damit betont sie mehr das Wie als das Was (bzw. Warum) sozialer Prozesse und macht bevorzugt auf die unhinterfragten Leistungen und Aktivitäten aufmerksam, die zur Hervorbringung sozialer Phänomene notwendig sind. Damit eröffnet sie die Möglichkeit, den Konflikt als eine eigenständige, von konkreten Seinsqualitäten unabhängige Form zu betrachten, die weder auf spezifische Eigenschaften und Motive noch auf besondere Kontexte zurückgreifen muss. Eine Konfliktanalyse, die sich ihren Gegenstand nicht von äußeren Einflüssen diktieren und vorgeben lässt, wird auf Konflikteigenschaften stoßen, die primär dem Gegenstand selbst und weniger seinen Umwelten zurechenbar sind. Eine konstruktivistische Konfliktforschung in dem hier vertretenen Sinne fordert deshalb mit Nachdruck dazu auf, ihre Einsichten so dicht als möglich am Gegenstand selbst zu entwickeln. Ihr analytischer Ausgangspunkt liegt in der realen Kommunikation begründet, die ihr empirisches Datum und Prüfstein gleichzeitig ist. Je näher sich die Analyse an dieser Prämisse orientiert, umso präziser lassen sich die analytischen Abstraktionsleistungen am Gegenstand selbst überprüfen.

# **Terrorismus-Berichterstattung und ihre Folgen**

Wolfgang Frindte, Nicole Haußecker & Jens Jirschitzka

"Journalists do not simply report on the world, they also assign meaning to the facts they report and interpret them according to a particular framework"

Kempf 2005, S. 3

### 1 Der Terrorismus als Inszenierung

Die öffentlichen Diskurse über den internationalen und transnationalen Terrorismus, die von ihm ausgehenden Gefahren und die Wege seiner Bekämpfung sind nach wie vor divers. Das betrifft die öffentlich geäußerten Meinungen über die Terrorrisiken, die Diskussionen über die Ursachen des Terrorismus und seine Akteure, die Auffassungen über die innen- und außenpolitischen Maßnahmen der Terrorbekämpfung im Allgemeinen und die Einstellungen zu den Sicherheitsstrategien der Bundesregierung und den Einsatz der Bundeswehr in Afghanistan im Besonderen. Dass die Protagonisten der Diskurse (die Politiker, die Journalisten, die wissenschaftlichen Experten, das Publikum) dabei unterschiedliche Auffassungen vertreten, dürfte gegenstandsimmanent sein. Terrorismus und die von ihm ausgehenden Gefahren sind Risiken, die eintreten könnten, aber keinesfalls mit Notwendigkeit eintreten müssen und deshalb mit einem hohen Maß an Irrealität verknüpft sind, wie Ulrich Beck meint:

"Risiken sind soziale Konstruktionen und Definitionen auf dem Hintergrund entsprechender Definitionsverhältnisse. Sie existieren in Form eines (wissenschaftlichen und alternativ-wissenschaftlichen) Wissens. Folglich kann ihre 'Realität' dramatisiert oder minimiert, verwandelt oder schlicht geleugnet werden gemäß den Normen, nach denen über Wissen und Nichtwissen entschieden wird. Sie sind Produkte von Definitionskämpfen und Definitionskonflikten im Rahmen bestimmter Definitionsmachtverhältnisse, also (mehr oder weniger erfolgreiche) Resultate von Inszenierungen" (Beck 2007, S. 66; Hervorhebung im Original).

Auch die Versuche, Terrorismus und Terror zu definieren, sind zahlreich und umstritten. Einige dieser Definitionen heben vor allem die Motive, andere dagegen vor allem die Ziele der terroristischen Aktionen (Weimann & Brosius 1989) hervor. In wieder anderen Begriffsbestimmungen steht die kriminelle Gewaltaffinität des Terrorismus im Vordergrund (z. B. die Definition der United Nations oder der Task Force on Disorder and Terrorism; vgl. Biernatzki 2002). Die intendierte Erzeugung psychischer Effekte (wie Angst, Furcht und Schock) steht im Mittelpunkt einer wei-

teren Gruppe von Definitionen (z. B. Brinkemper 2002, S. 212). Die Neutralität der Opfer bzw. die Unterscheidung zwischen Terrorzielen und Terroropfern bilden ebenfalls einen Kern möglicher Begriffsbestimmungen (Schmid & de Graaf 1982, S. 15). Hoffman (2002), Jenkins (1975), Tsfati & Weimann (2002), Waldmann (2005) u. a. betonen überdies, dass auch die angestrebte massenmediale Verbreitung der geplanten und/oder realisierten terroristischen Aktionen zu den Merkmalen von Terrorismus zu rechnen seien. Mediale Berichterstattung schaffe Öffentlichkeit und nur dadurch sei es den Terroristen überhaupt erst möglich, ihre Ziele zu realisieren. Um die in den vielen Definitionen angesprochenen Ziele des Terrorismus (Unsicherheit, Schrecken, Schadenfreude und Sympathie) erreichen zu können, bedarf es der medialen Inszenierung von Terror, Terroranschlägen und Terrorismus. Jenkins (1975) und Tsfati & Weimann (2002) sprechen deshalb vom "theater of terror". Und das führt uns zu folgender Arbeitsdefinition:

Terrorismus (von lat. *terror*: Furcht, Schrecken) ist 1. eine kalkuliert inszenierte gewalttätige Kommunikationsstrategie, mit der 2. (nichtstaatliche) Akteure versuchen, die Gesellschaft, Staaten, deren Institutionen oder bestimmte gesellschaftliche Gruppen zu schädigen und/oder 3. in Angst und Schrecken zu versetzen, um 4. auf diese Weise politische Ziele zu erreichen.

Moderner Terrorismus ist auf Verbreitungsmedien angewiesen und spekuliert auf deren Wirkungen. Seine effektive Inszenierung bemisst sich dabei nicht nur an der symbolhaften und spektakulären Gewaltperformance oder der medialen Aufführung durch Journalisten und Medienmacher. Vielmehr sind die Akteure und Sympathisanten des Terrorismus, die Ziele, die Opfer, die politischen, wissenschaftlichen und Alltagsbeobachter und die Medien gleichermaßen an der Inszenierung beteiligt. Nur so erhält der Terror seine Form und Wirkung.

# 2 Ziel- und Fragestellungen sowie theoretische Grundlagen

Die Frage lautet nicht, wie sich Nachrichten über Terrorereignisse in den Meinungen der Rezipienten widerspiegeln, sondern, ob und wenn ja, in welcher Weise die von Medien produzierten Konstruktionen über Terrorismus den individuellen Interpretationen der Medienrezipienten entsprechen, sie verstärken oder verändern können. Daraus leiten wir zwei zentrale Fragestellungen ab: 1. Welche Rolle spielen die Verbreitungsmedien in der Terrorismus-Inszenierung? 2. Wie nehmen Vertreter der deutschen Bevölkerung die Terrorismus-Inszenierung wahr und mit welchen Interpretationen beteiligen sie sich daran?

Wenn Inszenierungen Angebote für die soziale und individuelle Konstruktion von Wirklichkeit bereitstellen und wir nach den Formen oder Rahmen dieser Angebote fragen, so öffnet sich ein Forschungsfeld, das Jörg Matthes (2007, S. 150) das Framing-Forschungsprogramm nennt. Die wohl bekannteste operationale Definition von Frames, auf die sich – bei aller Divergenz – die meisten Protagonisten des Framing-Forschungsprogramms stützen, stammt von Robert M. Entman (1993, S. 52): "To frame is to select some aspects of a perceived reality and make them

more salient in a communicating text, in such a way as to promote a particular problem definition, causal interpretation, moral evaluation, and/or treatment recommendation".

In diesem Sinne werden durch Frames 1. bestimmte Sichtweisen oder Definitionen auf bzw. von Wirklichkeit hervorgehoben, 2. *Bewertungen* der dargestellten Wirklichkeit nahe gelegt, 3. *Erklärungen über Ursachen* für die so dargestellten Wirklichkeitsausschnitte bereit gestellt und 4. *Möglichkeiten zur Bewältigung* dieser Wirklichkeit offeriert. Frames liefern individuelle und/oder soziale (u. a. mediale) Interpretationsrahmen oder Bezugssysteme zur (Re-)Konstruktion von Wirklichkeit. Im Hinblick auf die mediale Konstruktion von Wirklichkeit und der damit angebotenen Interpretationsrahmen sprechen wir von Medienframes und unterscheiden formal-stilistische und inhaltsbezogene Frameelemente. Während sich formal-stilistische Elemente auf die Struktur und Präsentation (also auf das *Wie*) einer Medienbotschaft beziehen (z. B. auf die Beschaffenheiten der visuellen Darstellung), geht es bei den inhaltsbezogenen Frameelementen um das *Was* von Medienbotschaften, um die Darstellung und Rahmung der Medieninhalte.

Wir haben uns deshalb einerseits an der o. g. Definition von Entman (1993) orientiert und die von ihm genannten vier Aspekte medialer Frames (problem definition, causal interpretation, moral evaluation, and/or treatment recommendation) als inhaltsbezogene Frameelemente bestimmt und sie andererseits durch ein formal-stilistisches Element, das Ausmaß und die Effekte der medialen Dramatisierung, ergänzt (vgl. Abb. 1).



Abbildung 1: Elemente terrorrelevanter Medienframes

Die Terroranschläge am 11. September 2001 haben dazu geführt, dass viele Forschungsprojekte ins Leben gerufen wurden, die sich nicht nur mit den Stimmungen und Einstellungen in der Bevölkerung, sondern auch kritisch mit den Medieninhalten und ihren Gestaltungen beschäftigen. So wird der vornehmlich auf Ereignisse fokussierten Berichterstattung u. a. vorgeworfen, die Darstellung von Ursache-Wirkungs-Zusammenhängen und Konfliktanalysen zu vernachlässigen und statt-dessen emotionalisierende Merkmale und negative Stereotypisierungen in den Vordergrund zu rücken (vgl. u. a. Becker 2002, Werthes et al. 2002). Auch die Terrorismus-Berichterstattung in den Bildmedien scheint mit der ansteigenden Tendenz zu Visualisierung und Emotionalisierung eine immer größere Rolle in der Inszenierung des Terrorismus zu spielen (vgl. z. B. Maier 2003). Eine stark dramatisierende und emotionalisierende Berichterstattung nähert sich nicht nur den Inszenierungszielen der Terroristen an, sondern beeinflusst auch die Wahrnehmung, Bewertung

und Interpretation des Terrorismus in besonderer Weise und kann stereotype und vorurteilsbehaftete Sichtweisen darauf verstärken (vgl. auch Haußecker 2007).

Vor diesem theoretischen Hintergrund haben wir die erste zentrale Fragestellung (Rolle der Verbreitungsmedien in der Terrorismus-Inszenierung) folgendermaßen spezifiziert: 1. Welche Sichtweisen auf oder Definitionen von Terrorismus und Terroristen werden angeboten? 2. Wie werden Terrorismus und die von ihm ausgehende Gefahr dargestellt und bewertet? 3. Welche Erklärungsmuster über die Ursachen des transnationalen Terrorismus und seine Folgen liefern die Berichte? 4. Über welche Möglichkeiten zur Bewältigung des Terrorismus (z. B. Anti-Terror-Maßnahmen) wird wie berichtet? 5. Welche Dramatisierungsmittel lassen sich identifizieren?

In der Beantwortung der zweiten zentralen Fragestellung greifen wir die in der Literatur nicht unübliche Unterscheidung von Medienframes und Individual- bzw. Rezipientenframes auf (u. a. Kempf 2006a, Matthes 2007, Neuman et al. 1992, Scheufele 2003). Dabei nehmen wir an, dass Individualframes, mit denen Personen Terrorismus und terroristische Gefahren interpretieren, bewerten, erklären und zu bewältigen versuchen, eine analoge Struktur wie entsprechende Medienframes aufweisen können (vgl. Abb. 2).



Abbildung 2: Elemente terrorrelevanter Individualframes (in Anlehnung an Entman 1993)

Um diese Elemente von Individualframes gruppieren sich auch die wichtigsten Forschungsschwerpunkte, mit denen sich seit 2001 die sozialwissenschaftlichen Studien zum Terrorismus beschäftigen. Besonders dominant sind jene Untersuchungen, die sich auf die emotionalen Bewertungen, Befindlichkeiten und das Bedrohungserleben angesichts der Terrorrisiken beziehen (z. B. Carroll et al. 2006, Klein et al. 2009, u. v. a.). Eine besondere Rolle nehmen verstärkt seit etwa 2005 die Forschungen zum Anti-Terrorismus (val. z. B. Schneckener 2006) und zu den Einstellungen gegenüber Anti-Terror-Maßnahmen ein (vgl. z. B. Cohrs et al. 2005a, b). Zudem werden zunehmend Studien veröffentlicht, in denen nach Zusammenhängen zwischen dem Erleben von Terrorbedrohungen und der Zustimmung zu bzw. der Ablehnung von nationalen und internationalen Sicherheits- und Anti-Terror-Maßnahmen gefahndet bzw. nach den Folgen von Anti-Terror-Maßnahmen für muslimische Minderheiten in den europäischen Ländern gefragt wird (z. B. Huddy et al. 2007, Staub 2007). Dabei zeigen nationale wie internationale Befunde nach 2001 enge Zusammenhänge zwischen negativen bzw. ablehnenden Einstellungen gegenüber relevanten Fremdgruppen (vor allem Muslimen), der Zunahme sog. Ingroup-Favorisierungen (Ethnozentrismus, Nationalismus und Patriotismus) und den Einstellungen gegenüber Terrorismus im Allgemeinen und der Angst vor Terroranschlägen im Besonderen (z. B. Bonanno & Jost 2006, Davies et al. 2008). Außerdem werden das Bedrohungserleben, die gruppenbezogenen Einstellungen (Ingroup-Favorisierungen und Outgroup-Ablehnungen) sowie anti- und pro-terroristische Stellungnahmen (z. B. die Befürwortung oder die Ablehnung von Anti-Terror-Maßnahmen) durch verschiedene personale und soziale Variablen moderiert bzw. mediiert, etwa durch das Alter, das Geschlecht, den Bildungsstand, durch individuelle autoritäre Orientierungen oder übergreifende Werthaltungen (z. B. Crowson et al. 2006, Esses et al. 2002, Skitka et al. 2006, Van Zelst et al. 2003).

Diese vielfältigen und z. T. auch recht unterschiedlichen sozialwissenschaftlichen Befunde erlauben es, auch die zweite zentrale Fragestellung zu spezifizieren: 1. Wie werden Terrorismus und Terroristen individuell wahrgenommen und interpretiert? 2. Wie werden die Terrorismusgefahr und der "Kampf gegen den Terrorismus" bewertet und u. U. als bedrohlich erlebt? 3. Wie werden die Ursachen und die Folgen des Terrorismus beurteilt und interpretiert? 4. Inwiefern wird den Anti-Terror-Maßnahmen zugestimmt bzw. unter welchen Bedingungen und mit welchen Begründungen werden solche Maßnahmen abgelehnt? 5. Welche gruppenbezogenen Einstellungen (Ingroup-Favorisierungen, Outgroup-Ablehnungen) äu-Bern die Befragten und inwieweit hängen diese mit der Beurteilung des Terrorismus und des Anti-Terrorismus und mit dem Bedrohungserleben zusammen? 6. Gibt es Zusammenhänge zwischen der Beurteilung des Terrorismus und der Beurteilung von Anti-Terror-Maßnahmen, dem Bedrohungserleben und dem Mediennutzungsverhalten (z. B. den Medienpräferenzen und TV-Präferenzen, der Einschätzung, inwieweit soziale Bezugsgruppen die individuelle Mediennutzung beeinflussen)?

# 3 Methodische Konsequenzen

# 3.1 Medienanalyse zur Beantwortung der Fragestellung 1 (Rolle der Verbreitungsmedien)

Um die Rolle von Verbreitungsmedien in der Terrorismus-Inszenierung zu untersuchen, konzentrierten wir uns ausschließlich auf das Fernsehen, das noch immer das zentrale Leitmedium darstellt (Egger & Eimeren 2008, S. 581).

Für unsere Analyse wählten wir auf der Basis aktueller Zuschauerzahlen (vgl. Basisdaten Media Perspektiven 2006) folgende Abendnachrichten aus: ARD-Tagesschau, ZDF-heute, RTL-aktuell und Sat.1-Nachrichten. Die Nachrichten von ARTE und n-tv wurden gemeinsam mit den vier Hauptsendern für eine deskriptive Voranalyse herangezogen. Für die Hauptauswertung von ARD, ZDF, RTL und Sat.1 konnten im Zeitraum von August 2007 bis Februar 2009 insgesamt 1.145 terrorrelevante Fernsehbeiträge identifiziert, codiert und mit mehr als 15.000 Variablen beschrieben werden.

# 3.2 Panelstudie zur Beantwortung der Fragestellung 2 (Individuelle Interpretationen)

#### Operationalisierung terrorismusrelevanter Individualframes als individuelle Einstellungsmuster (halbstandardisiertes Interview)

#### Individuelle Einstellungsmuster

zur Interpretation, Bewertung, Erklärung und Bewältigung terrorrelevanter Wirklichkeiten

# Subjektive Interpretationen und (naive) Konzepte

über Terror, Terrorismus und Terrorakteure

#### Bewertungen von und Bedrohung durch Terrorismus

persönlich und/oder national relevant, als realistisch und/oder symbolisch

#### **Attributionen**

über Ursache, Wirkung und Wahrscheinlichkeit von Terror

#### Anti-Terrorismus

Akzeptanz anti-terroristischer Maßnahmen

#### Pro-Terrorismus

Akzeptanz terroristischer Gruppierungen und Aktionen

#### Operationalisierung möglicher Prädiktoren, Moderatoren bzw. Mediatoren der individuellen Einstellungsmuster (standardisierter Fragebogen)

#### Prädiktoren, Moderatoren bzw. Mediatoren der individuellen Einstellungsmuster

#### Soziodemografische Merkmale

Alter, Geschlecht, Bildung, Familienstand, Einkommen, Anzahl der Kinder

#### Generalisierte Einstellungen

Autoritarismus Wertorientierungen politische Orientierungen Religion, Religiosität

#### Ingroup-Favorisierung Nationalismus, Patriotismus

Outgroup-Ablehnung Ausländerfeindlichkeit, Antisemitismus, Ablehnung von Muslimen, Pro- bzw. Anti-Amerikanismus

#### Medienwahl- und Mediennutzungsverhalten

Dauer, Präferenzen, Motive, Beeinflussbarkeit (Dritte-Person-Effekt)

Abbildung 3: Themenschwerpunkte des standardisierten Fragebogens und des halbstandardisierten Interviews

Um individuelle Einstellungsmuster zur Interpretation, Bewertung, Erklärung von und zum handlungsbezogenen Umgang mit Terrorismus analysieren zu können, entschieden wir uns, eine Kombination verschiedener und aufeinander bezogener Erhebungs- und Auswertungsverfahren einzusetzen. Umgesetzt haben wir dies, indem 100 Personen mittels eines standardisierten Fragebogens und eines halbstan-

dardisierten Interviewleitfadens über drei Erhebungszeitpunkte (Welle 1: November bis Dezember 2007, Welle 2: Mai bis Juli 2008 und Welle 3: Januar bis Februar 2009), wiederholt interviewt, befragt bzw. um Dokumentation ihrer Mediennutzungsgewohnheiten gebeten wurden. Die Auswahl der Probanden richtete sich nach dem Alter, dem Geschlecht und der Bildung, wobei eine Gleichverteilung in Bezug auf diese Attribute angestrebt wurde. Mit dem standardisierten Fragebogen sollten relevante Einstellungen zum Terrorismus erfasst und operationalisiert werden. Mit dem Leitfaden für die halbstandardisierten Interviews wurden den Teilnehmern Fragen vorgelegt, die sie frei beantworten konnten. Abbildung 3 gibt einen Überblick über die angesprochenen Themenbereiche.

## 4 Ausgewählte Ergebnisse

# 4.1 Medienanalyse

Tabelle 1 gibt einen Überblick über die senderspezifische Verteilung der Beiträge und einen Einblick in die dominierenden Themen der Terrorismus-Berichterstattung.

Anzahl	ARD n	ZDF n	RTL n	Sat.1 n	Gesamt n
Beiträge über Terrorismus – Gesamt	368	323	273	181	1145
Beiträge, in denen über mindestens ein terroristisches Ereignisse berichtet wird	174	153	150	101	578
Beiträge über Terrorbedrohung	94	97	66	50	307
Beiträge über Ursachen des Terrorismus	41	48	34	33	156
Beiträge über Anti-Terror-Maßnahmen	239	205	151	112	707

Tabelle 1: Senderspezifische Verteilung der terrorismusrelevanten Beiträge

Augenscheinlich berichten die öffentlich-rechtlichen Sender ARD und ZDF insgesamt häufiger über Terrorismus als die privaten Sender RTL und Sat.1. Im Aggregat aller Sender war an fast jedem zweiten Tag (44% aller Publikationstage) ein Beitrag darüber zu sehen.

Zwei Themenkomplexe dominieren im Querschnitt in den untersuchten Sendern – der Kampf gegen Terrorismus und terroristische Ereignisse. Weitere Themen wie Terrorgefahren und Hintergründe sind dagegen weit seltener präsent. Die Sender scheinen also die Zuschauer kaum über das Wesen des Terrorismus und dessen Hintergründe aufzuklären, sondern vor allem ereignisbezogene Primärberichterstattung zu leisten.

Unsere quantitative Mediennanalyse der Fernsehberichterstattung (vgl. Früh 1998) bestätigt damit, was bereits andere Studien nach dem 11. September 2001

festgestellt und kritisiert haben (vgl. z. B. Biernatzki 2002, Werthes et al. 2002): Die Fernsehzuschauer müssen den Eindruck gewinnen, dass es vor allem auf den Kampf gegen Terrorismus ankommt, während Erklärungen und Ursachensuche eher nebensächlich sind (vgl. Abb. 4).

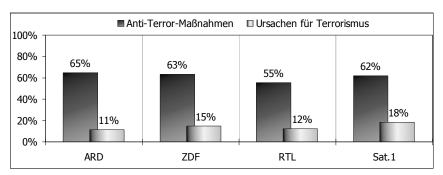
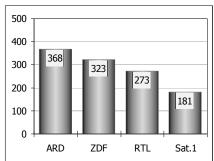


Abbildung 4: Ursachen und Maßnahmen in den Fernsehnachrichten von ARD, ZDF, RTL und Sat.1 – Prozentuale Anteile an der senderspezifischen Gesamtbeitragszahl als Basis (die Summe der Anteilswerte pro Sender ergibt in Folge der vorgenommenen Codierung nicht 100%)

Dabei dominieren vor allem Maßnahmen, die man unter dem Etikett Krieg/Kampf gegen Terrorismus zusammenfassen kann. Dazu gehören polizeiliche Maßnahmen, innenpolitische Sicherheitsmaßnamen, sowie militärische Maßnahmen und Aspekte der Außenpolitik. Entwicklungspolitische Maßnahmen thematisieren alle vier Sender kaum – und falls doch, dann fast immer im Kontext militärischer Maßnahmen.



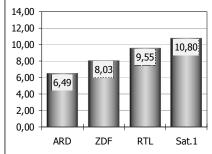


Abbildung 5: Anzahl der Beiträge über Terrorismus und durchschnittliche Kameraeinstellungen pro Beitrag in den Fernsehnachrichten von ARD, ZDF, RTL und Sat.1 – absolute Häufigkeiten Die linke Hälfte zeigt die absoluten Häufigkeiten der Beiträge über Terrorismus pro Sender, während die rechte Hälfte die mittlere Anzahl der Kameraeinstellungen pro Beitrag (Mittelwerte) wiedergibt.

Wie wir weiter oben ausgeführt haben, können Visualisierung und Emotionalisierung entsprechende mediale Bezugsrahmen unterstützen und verstärken. Abbildung 5 stellt nach Sendern die Anzahl themenrelevanter Nachrichtenbeiträge (linke Hälfte) der Anzahl der Visualisierungen (rechte Hälfte) gegenüber. Daran wird

deutlich, dass die öffentlich-rechtlichen Sender zwar häufiger als die Privatsender über Terrorismus berichten, dabei aber visuell zurückhaltender sind. Dagegen hat Sat.1 die wenigsten Beiträge, jedoch die mit den meisten Einstellungswechseln. Anders ausgedrückt lässt sich festhalten, dass die Beiträge der Privatsender mehr terrorismusrelevante Visualisierungen bzw. eine schnellere Schnittfrequenz (cuts) aufweisen, was ihnen eine gewisse Dynamik verleiht. Solch "schnelle Schnitte" (Bonfadelli 1993, S. 166) können bei den Zuschauern eine physiologische Erregung bzw. entsprechende Emotionen hervorrufen, wie dies auch andere Gestaltungsmerkmale wie z. B. Nahaufnahmen oder ungewöhnliche Perspektiven tun.

Laut Maier & Stengel (2007) hat die Visualisierung in den letzten zehn Jahren in der Berichterstattung über Krieg und Terrorismus deutlich zugenommen. Die dramatischen und häufig wiederholten Bilder der einstürzenden Türme des WTC vom 11. September sind dafür prototypisch. Diese Visualisierungen sind letztlich genau das, worauf es die Terroristen angelegt haben – auf weltweit ausgestrahlte Bilder voller Dramatik und Emotion. Waldmann (2005) spricht in diesem Kontext von einer einseitigen, verzerrten Mediendarstellung. Ob auch die von uns untersuchten Fernsehnachrichten von Dramatisierung und Emotionalisierung Gebrauch machen, haben wir anhand unterschiedlicher Indikatoren untersucht. Abbildung 6 zeigt, dass RTL und Sat.1 häufiger als ARD und ZDF auch Bilder von Opfern zeigen.

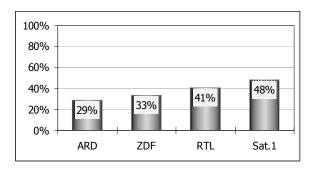


Abbildung 6: Visualisierung von Opfern in den Fernsehnachrichten von ARD, ZDF, RTL und Sat.1 – Anteile an allen Beiträgen

Nun erweitern wir unsere Analysen und suchen nach übergeordneten Dimensionen bzw. Faktoren, die die Terrorberichterstattung über alle Beiträge auszeichnen und welche die Rezipienten in ihrer Sicht auf den Terrorismus u. U. beeinflussen könnten. Wir nehmen an, dass mögliche Wirkungen der Terrorismus-Berichterstattung nicht von einzelnen Beiträgen abhängig sind, sondern von charakteristischen Themensetzungen, über die Beiträge und u. U. auch über die Sender hinweg.

In einem ersten Schritt wird mit einer Faktorenanalyse versucht, die korrelativen Zusammenhänge zwischen mehreren manifesten Einzelvariablen (in diesem Falle die aggregierten Kategorienvariablen) mit einer kleineren Anzahl zugrundeliegender Faktoren bzw. Dimensionen zu erklären.

Tabelle 2 zeigt die MSA-Werte, die Kommunalitäten (Anteil der durch die Faktoren aufgeklärten Varianz einer Variablen) und die standardisierten Regressionskoeffizienten der Oblimin-Faktoren für die einzelnen Variablen.

Variablen		KM	Part	Partielle Regressionskoeffizienten de						er Fak	toren
			1	2	3	4	5	6	7	8	9
Thematisierung von Gefühlen	.81	.61	.75								
Visuelle Authentizität		.77	.60								
Visuelle Opferdarstellung		.58	.65								
Dramatisierung in Sprache/Ton		.51	.68								
Polizei im Einsatz (visualisiert)		.42	.40								
Aktualität (Zeitfokus)	.70	.73		.86							
Terroranschlag	.72	.73		.72							
Trümmer/Zerstörung (visuell)	.84	.63		.62							
Zivilisten als Opfer	.74	.51		.65							
Religion der Täter: Islam	.80	.85			.91						
Religiöser Terrorismus	.79	.84			.90						
Religiöse Motive der Täter	.81	.84			.90						
Bedrohung Deutschlands	.79	.58				.67					
Unbestimmte Terrorgefahr	.79	.50				.64					
Allgegenwart der Gefahr		.51				.59					
Polizei im Innendienst (visuell)	.69	.48					.69				
Sicherheitspolitik im Innern	.62	.43					.59				
Polizeiliche Maßnahmen	.79	.49					.52				
Justiz (visualisiert)		.63						.79			
Juristische Maßnahmen	.58	.69						.78			
Militär (visualisiert)	.58	.68							.81		
Militärische Maßnahmen	.61	.67							.80		
Entwicklungspol. Maßnahmen	.60	.34							.46		
Ethnisch-National. Terrorismus	.51	.73								.85	
Herkunft der Täter: Europa	.59	.67								.81	
Politische Tätermotive	.51	.41								.55	
Sozialrevolutionär. Terrorismus	.60	.81									.91
Extremistische Motive der Täter		.71									.83
Herkunft der Täter: Deutschland	.73	.65									.64
Politiker als Opfer	.66	.48									.54

Tabelle 2: MSA-Werte, Kommunalitäten (KM) und partielle Regressionskoeffizienten der Faktorgefügematrix nach Hauptkomponentenanalyse und Oblimin-Rotation (n = 1.145)

*Anmerkungen*: Die Tabelle berücksichtigt nur partielle Regressionskoeffizienten, die mindestens einen Wert von 0.40 aufweisen. Die Regressionskoeffizienten können größere Werte als 1.0 annehmen, aus Platzgründen wurde jedoch die für Korrelationen übliche Schreibweise verwendet. Die Faktoren 3, 6 und 7 wurden zur Vereinfachung der Interpretation umgepolt.

Nach dem Auffinden der Dimensionen wurden die Beiträge je nach deren Ausprägung mit clusteranalytischen Verfahren in Gruppen eingeteilt. Vor der Anwendung dieser Verfahren haben wir allerdings Einschränkungen vorgenommen: Bei Faktor 1 (Gestalterische Dramatisierung) handelt es sich um eine Dimension der formalen Gestaltung eines Nachrichtenbeitrages. Die anderen Dimensionen verweisen demgegenüber auf die inhaltliche Ausrichtung möglicher Medienframes. Zur inhaltlichen Gruppierung haben wir die formal-stilistische Dimension (Faktor 1) zunächst nicht in die nachfolgend beschriebenen clusteranalytischen Verfahren aufgenommen. Stattdessen haben wir separat über alle Beiträge aller vier Sender aus den Indikatormerkmalen von Faktor 1 ein Dramatisierungsindex gebildet, indem die zugehörigen Indikatorvariablen zunächst z-transformiert (d. h. auf einen Mittelwert von 0 und eine Streuung von 1 gebracht) und dann gemittelt wurden, dessen Reliabilität .73 (Cronbachs  $\alpha$ ) beträgt. Auf diesen Dramatisierungsindex kommen wir nach erfolgter Beitragsgruppierung wieder zurück.

In die anschließende Clusteranalyse mit Ward-Methode gingen 1.095 Nachrichtenbeiträge ein. Als Eingangsvariablen nutzen wir die Faktorwerte der Beiträge auf den verbleibenden sieben Faktoren, die wir als charakteristische Dimensionen der Terrorismus-Berichterstattung identifiziert hatten und die nun als Merkmale dienen, welche bei jedem einzelnen Beitrag mehr oder weniger stark ausgeprägt sein können.<sup>2</sup>

Folgende Cluster konnten identifiziert werden: Aktueller Terroranschlag, Islamistischer Terrorismus, Islamistische/Unbestimmte Terrorgefahr in Deutschland, Polizei und Sicherheitspolitik im Innern, Justiz und juristische Maßnahmen, Militär und militärische Maßnahmen, Ethnisch-Nationalistischer/Separatistischer Terrorismus und Terrorismus unspezifisch.

Abbildung 7 zeigt diese acht Themenschwerpunkte im Sendervergleich. Hier bestätigt sich weitgehend, dass die öffentlich-rechtlichen Sender mehr über Terrorismus berichten als die Privatsender; mit der Ausnahme im Themenschwerpunkt Islamistische/Unbestimmte Terrorgefahr für Deutschland, in dem öffentlich-rechtliche und private Sender in etwa gleich auf liegen.

Somit lässt sich feststellen: Potenzielle Rezipienten werden, sofern sie sich für bestimmte Gruppen von Nachrichten interessieren, nur ausschnitthaft mit den charakteristischen Dimensionen oder Mustern der Terrorismus-Berichterstattung konfrontiert.

<sup>1</sup> Eine zweite Einschränkung in Vorbereitung auf die clusteranalytische Auswertung betrifft Faktor 9 (Sozialrevolutionärer Terrorismus). Diese Dimension und die zugehörigen Indikatormerkmale weisen augenscheinlich auf den RAF-Terrorismus der 1970er-Jahre hin. Da unsere zentralen Forschungsinteressen auf die mediale Konstruktion und die individuellen Interpretationen insbesondere des transnationalen und aktuell bedeutsamen Terrorismus gerichtet sind, haben wir Faktor 9 ebenfalls nicht in die clusteranalytischen Gruppierungsverfahren einbezogen.

<sup>2</sup> Als Anhaltspunkt zur Bestimmung einer sinnvollen Clusteranzahl wurde das Struktogramm herangezogen (z. B. Bortz 2005). Ein auffälliger Sprung im Anwachsen der Fehlerquadratsumme zeigte sich zwischen den Fusionsstufen mit acht bzw. sieben Clustern, so dass die Lösung mit acht Clustern ausgewählt wurde.

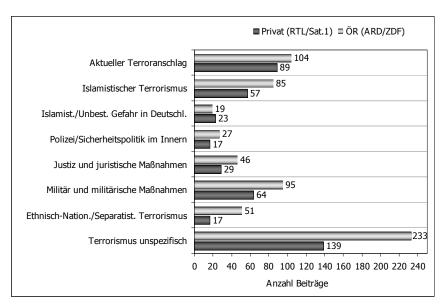


Abbildung 7: Sendergruppenspezifische Häufigkeitsverteilung der Beiträge von ARD/ZDF (n = 660) und RTL/Sat.1 (n = 435) auf die Beitragscluster – absolute Häufigkeiten (Gesamtbeitragszahl n = 1.095)

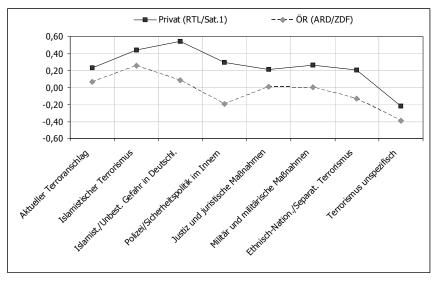


Abbildung 8: Mittlerer Dramatisierungsindex aller terrorismusbezogenen Beiträge (n = 1.095) der Sender ARD, ZDF, RTL und Sat.1 im Zeitverlauf – senderspezifische Mittelwerte pro Monat

Wenn es – abgesehen von der höheren Beitragsanzahl der öffentlich-rechtlichen Sender – in den inhaltlichen Dimensionen keine auffallenden Unterschiede zwischen den Sendern gibt, ist zu fragen, ob sich in den formal-stilistischen Dimensionen, also in der Gestaltung der Berichterstattung, senderspezifische Unterschiede nachweisen lassen.

Um diese Frage beantworten zu können, haben wir für alle analysierten Nachrichtenbeiträge einen Dramatisierungsindex berechnet, der aus folgenden Variablen gebildet wurde: Visuelle Opferdarstellung, Thematisierung von Gefühlen, visuelle Authentizität, Dramatisierung in Sprache und Ton und visualisierter Polizeieinsatz. In Abbildung 8 sind die senderspezifischen Dramatisierungsmittelwerte für die o. g. acht Themenschwerpunkte veranschaulicht.

Die Unterschiede zwischen den öffentlich-rechtlichen und den privaten Sendern sind statistisch signifikant (p < .001), d. h. die Beiträge der Privatsender weisen unabhängig von den Themenschwerpunkten einen im Durchschnitt höheren Dramatisierungsgehalt auf.

Diese Unterschiede zeigen sich auch im Zeitverlauf. Statistische Prüfungen (mittels Mann-Whitney-U-Tests) bestätigen sie und verweisen auf signifikante rangplatzbezogene Unterschiede bzgl. des Dramatisierungsindex zwischen Öffentlich-Rechtlichen und Privaten in den Monaten September und Oktober des Jahres 2007 und in den Monaten Februar, März, September, November und Dezember des Jahres 2008.

Besondere Dramatisierungsspitzen finden sich bei den Privatsendern in den Monaten vor den Panelbefragungen und -interviews, die wir zur Beantwortung der Fragestellung 2 unseres Projekts durchgeführt haben. Die Befragungen und Interviews fanden zwischen November 2007 und Januar 2008, Mai und Juli 2008 sowie Januar und März 2009 statt. Die auffallenden Dramatisierungsspitzen der Privatsender liegen im September und Oktober 2007, im Februar und März 2008 sowie im September, November und Dezember 2008.

Geht man nun von der Annahme aus, dass potenzielle Nachrichtenrezipienten aus der Panelstudie (aufgrund ihrer Senderpräferenzen) vor allem Privatsender nutzen, um sich über den Terrorismus zu informieren, dann liegt die Vermutung nahe, dass sie im Vorfeld der Befragungen und Interviews mit Nachrichten konfrontiert wurden, die einen besonders hohen Dramatisierungsgehalt besitzen. Welche Folgen eine solche mediale Konfrontation haben könnte, werden wir im nächsten Abschnitt darstellen.

# 4.2 Ausgewählte Ergebnisse aus den Interviews und Befragungen

Im halbstandardisierten Interview fragten wir u. a. "Was stellt Ihrer Meinung nach derzeit die größten weltweiten Bedrohungen dar?" und gaben die Möglichkeit vor, mehrere Bedrohungen zu nennen. Abbildung 9 zeigt nur die Bedrohungsarten, die von den Interviewten in der ersten Erhebungswelle (Ende 2007–Anfang 2008) auf

den ersten Rangplatz gesetzt wurden. Auf die Frage antworteten nur 88 Probanden, von denen fast die Hälfte (43) der Meinung war, die größte Bedrohung gehe momentan vom Klima und von der Umwelt aus. Nur 17 Interviewpartner sahen im Terror und im Terrorismus die größte Bedrohung.

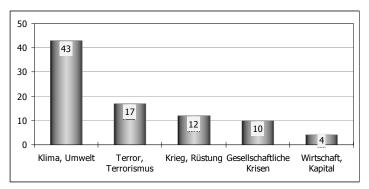


Abbildung 9: Erstgenannte größte Bedrohung weltweit, Welle 1 (n = 100, geantwortet n = 88)

Wir betrachten im nächsten Schritt die entsprechenden Befunde aus der standardisierten Befragung. Im Fragebogen wurde das persönliche Erleben von Bedrohung durch den Terrorismus mit der sehr reliablen Skala *Erleben persönlicher Bedrohung* erhoben (Items: "Fühlen Sie sich persönlich durch Terroranschläge bedroht?", "Haben Sie Angst vor Terroranschlägen hier in Deutschland?", "Sind Sie besorgt, ob Sie in Deutschland vor Terroranschlägen sicher sind?"). Die deskriptive und prüfstatistische Auswertung der Antworten zeigt folgende Ergebnisse:

- Geschlechterunterschiede: Nur in der Erhebungswelle 1 fühlten sich Frauen durch den Terrorismus stärker persönlich bedroht (M = 2.76) als Männer (M = 2.13; t(97) = 2.40, p < .01). In den Wellen 2 und 3 zeigten sich keine Geschlechterunterschiede.
- Altersunterschiede: In Welle 1 äußerten die jüngeren Probanden ein signifikant geringeres persönliches Bedrohungserleben (M = 2.07) als die älteren (M = 3.70; t(74) = -2.26, p < .05). Dies trifft auch für Erhebungswelle 3 zu (Jüngere M = 1.88; t(45) = -2.47; p < .05; Ältere: M = 2.70).</li>
- *Bildungsunterschiede:* Hinsichtlich des Bildungsgrades der Probanden ergaben sich über alle drei Erhebungswellen hinweg keine nennenswerten Unterschiede.
- Familienstand: In Erhebungswelle 1 fühlten sich verheiratete Personen stärker bedroht (M = 3.74) als ledige (M = 3.10, t(92) = -2.15, p < .05); was sich auch in Welle 2 zeigte (Verheiratete: M = 2.75 Ledige: M = 2.14, t(73) = -2.24, p < .05).</li>

Auch im Panelvergleich, also über alle drei Erhebungswellen hinweg, fühlen sich die Probanden, die an allen drei Wellen teilnahmen, im Jahre 2009 weniger bedroht als 2007 und 2008. Während sich Ende 2007 noch etwa 12% der Befragten mehr oder weniger persönlich vom weltweiten Terrorismus bedroht fühlten, waren es 2008 nur

9% und 2009 gar nur noch 6%. Die Unterschiede zwischen Welle 1 ( $M_1 = 2.63$ ) und Welle 3 ( $M_3 = 2.22$ ) sind signifikant (t(48) = 3.83, p < .001), ebenso die zwischen Welle 2 ( $M_2 = 2.45$ ) und Welle 3 ( $M_3 = 2.22$ ; t(48) = 2.32, p < .05). Das heißt: Zu Beginn unserer Studie, in der ersten Erhebungswelle zeigten sich sowohl in den Interviews als auch bei den Befragungen deutliche Unterschiede im Bedrohungserleben zwischen Männern und Frauen: Frauen fühlten sich bedrohter.

Das Ausmaß der erlebten Bedrohung angesichts von Terrorgefahren ist in unserer Stichprobe (über alle drei Erhebungswellen) wesentlich geringer als in der Gesamtbevölkerung. Das mag an der Besonderheit unserer Stichprobe liegen, könnte aber auch mit der sehr sensiblen Erhebungs- und Interviewmethode zusammenhängen. Relativ kompatibel mit den Befunden repräsentativer Meinungsumfragen sind dagegen unsere Panelergebnisse. Sowohl das nationale als auch das persönliche Bedrohungserleben hat sich über unsere drei Erhebungswellen signifikant verringert. Das ist einerseits nicht sonderlich überraschend. Andererseits scheinen diese und ähnliche Befunde im Widerspruch zu offiziellen Terrorwarnungen zu stehen, die in den Jahren 2008 und 2009 auch in Deutschland ausgegeben wurden.

Betrachtet man die Zustimmung zu Anti-Terror-Maßnahmen, so zeigte sich in der standardisierten Befragung, dass die Interviewten der Welle 1 signifikant eher solchen Maßnahmen zustimmen, die sich auf verantwortungsvolle Wirtschafts- und Entwicklungspolitik, auf humanitäre Hilfe und auf die Suche nach Gesprächen mit den Terroristen beziehen, als verstärkten Sicherheits- und Überwachungsmaßnahmen und militärischen Einsätzen. Gleiches gilt für Welle 2 und 3. Die Unterschiede in der Zustimmung zu den verschiedenen Anti-Terror-Maßnahmen sind in allen drei Erhebungswellen signifikant (p < .05) und damit statistisch bedeutsam. Wir verzichten auf die vollständige Darstellung der prüfstatistischen Ergebnisse und illustrieren die Befunde stattdessen und beispielhaft an der Erhebungswelle 1 (vgl. Abb. 10).

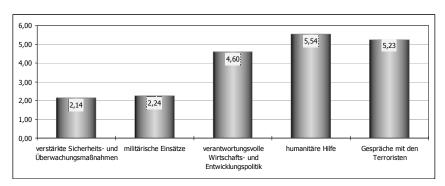


Abbildung 10: Zustimmung zu Anti-Terror-Maßnahmen – Welle 1 (Mittelwerte und Standardabweichungen)

Die Mehrheit der Befragten steht militärischen Anti-Terror-Maßnahmen, aber auch

<sup>3</sup> Die tiefgestellten Indizes bei den Angaben der Mittelwerte M<sub>1</sub>, M<sub>2</sub> und M<sub>3</sub> verweisen hier und im Folgenden auf die Erhebungswellen 1, 2 und 3.

verstärkten Sicherheits- und Überwachungsmaßnahmen eher skeptisch gegenüber.

Da wir die Probanden über zwei Jahre hinweg insgesamt drei Mal befragen und interviewen konnten, verfügen wir über ein Datensample, dass es uns auch erlaubt, auf statistischem Wege nach Wirkungszusammenhängen (Kausalitäten) zwischen einzelnen Variablen zu suchen. Aus Platzgründen verzichten wir auf die ausführliche Erläuterung der eingesetzten statistischen Methoden (Cross-Lagged-Regression-Analysen: vgl. ausführlich über das weitere Vorgehen Reinders 2006). Ein statistischer Hinweis auf Kausalität zwischen einer unabhängigen (in unserem Falle der Outgroup-Ablehnung bzw. der Ablehnung von Fremden) und einer abhängigen Variable (den Einstellungen gegenüber Anti-Terror-Maßnahmen) liegt dann vor, wenn erstens die Ursache der Wirkung zeitlich vorgelagert ist. Das heißt, die Messung der unabhängigen Variablen muss vor der Messung der abhängigen erfolgen. Zweitens müssen Ursache und Wirkung kovariieren. Es muss also ein nicht zufälliger Zusammenhang zwischen beiden Variablen nachgewiesen werden. Schließlich muss die Ursache die alleinige (oder mindestens hauptsächliche) Erklärung für die Wirkung darstellen. Die Effekte der unabhängigen auf die abhängige Variable müssen signifikant größer als andere Effekte ausfallen. Prinzipiell ist in diesem Sinne die Überprüfung von Kausalität nicht nur in Prä-Post-Experimenten möglich, sondern auch durch Panelstudien, so wie wir sie durchgeführt haben. Dazu müssen z. B. bei einer Befragung mit mindestens zwei Messzeitpunkten die unabhängigen und abhängigen Variablen zeitversetzt erfasst und deren empirische Relation zueinander ermittelt werden. Das haben wir durch die drei Erhebungswellen realisieren können. Das auf diesen Annahmen aufbauende Vorgehen in Cross-Lagged-Analysen ist in den Abbildungen 11-13 illustriert.

Nationale und internationale Befunde legten es zunächst nahe, nach kausalen Zusammenhängen zwischen Bedrohungserleben und der Zustimmung zu Anti-Terror-Maßnahmen zu fahnden. Wir prüften dies über die drei Wellen, fanden jedoch keine Effekte: Weder die Zustimmung zu noch die Ablehnung von Anti-Terror-Maßnahmen werden davon beeinflusst, ob das Bedrohungserleben gering oder stark ausgeprägt ist. Das heißt mit anderen Worten auch: Wenn Personen bestimmte Anti-Terror-Maßnahmen befürworten oder ablehnen, so tun sie dies zumindest nicht primär infolge ihres verstärkten Bedrohungsempfindens angesichts des internationalen Terrorismus.

Die Frage, die wir uns angesichts dieser Befunde stellten, lautete deshalb: Wann stimmen dann aber unsere Probanden gewissen Anti-Terror-Maßnahmen zu bzw. wann lehnen sie diese ab? In der weiteren statistischen Auswertung unserer Interviewdaten zeigten sich folgende Zusammenhänge, die durchaus als Wirkzusammenhänge zu interpretieren sind, auch wenn sie sich zunächst nicht so einfach erschließen (vgl. Abb. 11).<sup>4</sup>

Die Pfeile und die beigefügten Zahlen verweisen auf die kausalen Zusammenhänge zwischen den Variablen Outgroup-Ablehnung und Zustimmung zu militärischen Maßnahmen bzw. zu verstärkten Sicherheits- und Überwachungsmaßnahmen. Die Zahlen, die in der Regel zwischen –1 und +1 variieren, zeigen die Stärke der Zusammenhänge an. Werte nahe 1 drücken starke, Werte nahe 0 sehr schwache Zusammenhänge aus. Die Sternchen an den Zahlen geben an, dass die Zusammenhänge mit einer Irrtumswahrscheinlichkeit von 5%, 1% bzw. 0.1% signifikant bzw. hoch signifikant sind.

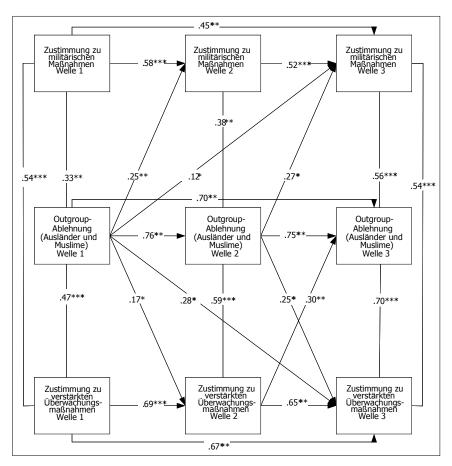


Abbildung 11: Querschnittliche, längsschnittliche und kausale Zusammenhänge zwischen Ablehnung fremder Gruppen (Ausländer und Muslime) und der Zustimmung zu Anti-Terror-Maßnahmen (+ p < .10, \*p < .05, \*\*p < .01, \*\*\* p < .001)

Die kausalen Pfade – die Pfade, die über einen Zeitpunkt hinaus reichen – zeigen, dass es über die Zeit unidirektionale kausale Wirkungen von ablehnenden Einstellungen gegenüber Fremdgruppen (Outgroups) zu Anti-Terror-Maßnahmen gibt, während die umgekehrten Kausalpfade nicht signifikant sind. Mit anderen Worten: Personen, die fremde Gruppen stärker ablehnen, befürworten deshalb auch verschärfte Anti-Terror-Maßnahmen, in diesem Falle militärische Einsätze und verstärkte Sicherheits- und Überwachungsmaßnahmen.

Die Ablehnung von Fremden bzw. die Outgroup-Ablehnung hat zwei Aspekte: die Ablehnung von Ausländern im Allgemeinen und die Ablehnung von Muslimen im Speziellen. Das heißt, Personen, die Ausländer *und* Muslime ablehnen, befürworten auch eher verschärfte Anti-Terror-Maßnahmen.

Wir haben des Weiteren gefragt, ob die Ablehnung von Ausländern *und* Muslimen mit dem Erleben von Terrorbedrohung zusammenhängt und fanden zunächst keine interpretierbaren Zusammenhänge. Deshalb haben wir das Gesamtkonstrukt Outgroup-Ablehnung wieder in seine o. g. Aspekte zerlegt und nach Prädiktoren für diese gesucht.

In statistischen Kausalanalysen über alle Interviewwellen hinweg zeigte sich, dass nur die Ablehnung von Muslimen in Welle 3 offenbar eine Folge der subjektiv wahrgenommenen persönlichen Terrorbedrohungen in den Wellen 1 und 2 ist. Je stärker sich Personen in Welle 1 und/oder 2 persönlich bedroht fühlten, desto stärker war ihre Ablehnung von Muslimen in Welle 3. Für den Aspekt Ablehnung von Ausländern wie auch für das Gesamtkonstrukt Outgroup-Ablehnung konnten diese Zusammenhänge nicht nachgewiesen werden. Abbildung 12 fasst die Befunde – ebenfalls ohne die nichtsignifikanten Zusammenhänge – zusammen.

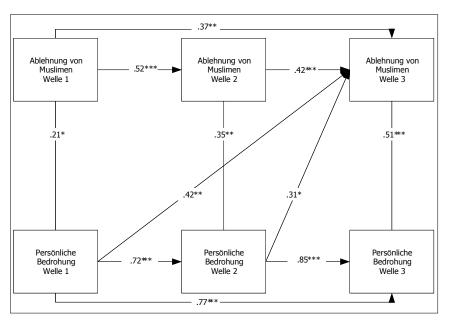


Abbildung 12: Kausale Zusammenhänge zwischen persönlichem Bedrohungserleben und der Ablehnung von Muslimen

(\* p < 
$$.05$$
, \*\* p <  $.01$ , \*\*\* p <  $.001$ ).

Eine Frage lag nun nahe: Erlauben unsere Daten Rückschlüsse darauf, ob und wenn ja, inwieweit der empirisch nachgewiesene Zusammenhang zwischen dem persönlichen Bedrohungserleben angesichts terroristischer Gefahren, der darauf aufbauenden Ablehnung von Muslimen und der verstärkten Zustimmung zu militärischen und sicherheitspolitischen Anti-Terror-Maßnahmen in Folge von Outgroup-Ablehnung eine Parallele zu den öffentlichen und medial verbreiteten Dis-

kursen über die vermeintliche Verknüpfung von Terrorismus, Islam und Muslimen aufweist?

Wir prüften mittels Cross-Lagged-Analysen, ob die Ablehnung von Muslimen von einer der im Panel operationalisierten Mediennutzungsvariablen (Nutzung von ARD, ZDF, RTL, etc.) kausal beeinflusst wird. Der einzige bedeutsame Effekt ist der Einfluss der RTL-Nutzung in Welle 2 auf die spezifische Ablehnung von Muslimen in Welle 3. Das heißt, Personen, die häufig RTL nutzen, um sich zu informieren, lehnen Muslime stärker ab (vgl. Abb. 13).

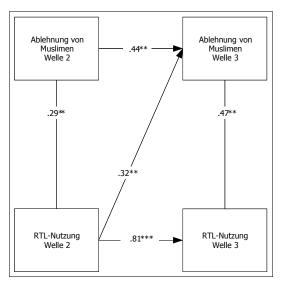


Abbildung 13: Kausale Zusammenhänge zwischen RTL-Nutzung und Ablehnung von Muslimen (\* p < .05, \*\* p < .01, \*\*\* p < .001)

Fragt man nun, wer die RTL-Nutzer sind, so zeigen statistische Vergleiche (mittels nichtparametrischem Mann-Whitney-U-Test) folgende signifikante Unterschiede zu Nicht-oder Kaum-RTL-Nutzern: RTL-Nutzer verorten sich eher im rechten politischen Spektrum (p < .01), sie sind nationalistischer eingestellt (p < .05), ausländerfeindlicher (p < .001), konformistischer (p < .01) und generell autoritärer orientiert (p < .000).

#### 5 Fazit

Sowohl in den öffentlich-rechtlichen als auch in den privaten Fernsehnachrichten wird über die Ursachen von Terrorismus relativ wenig berichtet; Anti-Terror-Maßnahmen dagegen werden umso häufiger thematisiert. Den Fernsehzuschauern werden also die Gefahren der Terrorbedrohung und die Notwendigkeit des Anti-Terrorkampfes vermittelt; Erklärungen der Terrorgefahr und ihrer Ursachen sowie

Begründungen für politisch beschlossene Anti-Terror-Maßnahmen hingegen werden kaum angeboten. Die Privatsender berichten zwar tendenziell weniger über Terrorismus, aber wenn sie dies tun, dann nutzen sie ausgeprägte Dramatisierungsstrategien stärker als die öffentlich-rechtlichen Fernsehsender. Das heißt, die Wirkung der Terrorismus-Berichterstattung scheint eher von deren Qualität als von der Quantität abzuhängen. Mehr als die Hälfte der Beiträge mit starken Dramatisierungseffekten stammen von den Privatsendern: Dabei handelt es sich vor allem um bildliche Darstellungen von Opfern, Verletzten/Toten oder um Berichte über mögliche Folgen vereitelter Terroranschläge mit dramatischen Sprach- und Toneffekten.

Die Interview- und Befragungsdaten der Panelstudie zeigen: Die Mehrheit der Befragten lehnt – über den gesamten Untersuchungszeitraum hinweg – verschärfte Sicherheits- und Überwachungsmaßnahmen und verstärkte Militäreinsätze im "Kampf gegen den Terrorismus" ab. Die in internationalen Studien nachgewiesenen Zusammenhänge zwischen dem persönlichen Bedrohungserleben und den Einstellungen gegenüber Anti-Terror-Maßnahmen konnten nicht bestätigt werden. In unserer Erhebung sind es vor allem jene Personen, die Muslime ablehnen sowie militärische Einsätze und verschärfte Sicherheits- und Überwachungsmaßnahmen im "Kampf gegen den Terrorismus" befürworten. Überdies zeigt sich, dass persönliches Bedrohungserleben angesichts terroristischer Gefahren die Ablehnung von Muslimen erhöhen kann. Das Bedrohungserleben wiederum hängt mit dem starken Einfluss der privaten Fernsehsender zusammen: Personen, die mehr Privatfernsehen als Öffentlich-Rechtliches schauen und welche folglich auch in ihrer Meinungsbildung über den Terrorismus vornehmlich von privaten Fernsehnachrichten beeinflusst werden, erleben stärkere persönliche Bedrohungen als diejenigen, die sich überwiegend durch die Öffentlich-Rechtlichen beeinflusst sehen oder möglicherweise auch gar nicht fernsehen. Andererseits lässt sich auch vermuten, dass sich die Nutzer des Privatfernsehens in ihren Einstellungen, Emotionen und Wertvorstellungen durch dieses bestätigt sehen und es deshalb präferieren.

Wir führen diese Effekte vor allem auf die stärker ausgeprägten Dramatisierungsstrategien der privaten Fernsehsender zurück und schlussfolgern: Nachgewiesenermaßen gehen vom transnationalen und internationalen Terrorismus lokale und globale Gefahren aus; der Umgang damit wird allerdings nicht leichter, wenn die Terrorgefahren und -risiken – vor allem von den privaten Fernsehsendern – in medial inszenierter Weise dramatisiert werden.

Somit könnte man eine Kaskade individueller Interpretationen des Terrorismus postulieren (vgl. Abb. 14).

Jean Baudrillard (2003, S. 32) formulierte zugespitzt: "Es gibt keine gute Weise des Mediengebrauchs, die Medien sind Teil des Ereignisses, sie sind Teil des Terrors, und sie wirken im einen oder im anderen Sinne". Zweifellos treibt er damit die Medienkritik auf die Spitze. Den Massenmedien kann zwar einerseits kaum die Schuld oder die Verantwortung für die Terroranschläge der letzten Jahre zugewiesen werden; sie schaffen aber andererseits die Voraussetzungen dafür, dass lokale

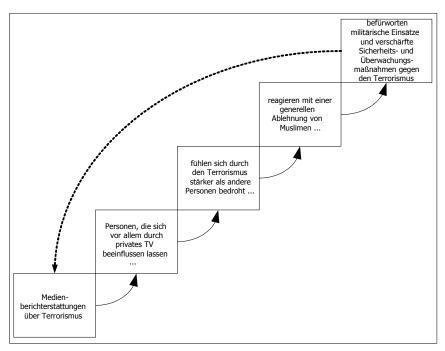


Abbildung 14: Kaskade der Wirkung von der Medienberichterstattung über den Terrorismus auf die individuellen Interpretationen über den Terrorismus, seine psychologischen Folgen und Konsequenzen

terroristische Ereignisse globale Wirkungen erzielen und ein globales Publikum finden. Moderner Terrorismus ist auf die Funktion der Verbreitungsmedien angewiesen und spekuliert auf deren Wirkungen. Das ist dann auch die Mahnung, die wir den Medienmachern ins Stammbuch schreiben möchten.

# "Worte, die wehtun, und Bilder, die schrei'n." Friedenskitsch am Beispiel von Songs im Eurovision Song Contest

Susanne Jaeger & Michael Reimann

"Der Widerspruch zwischen maßloser Medienpräsenz und gleichzeitig lückenloser Zensur während des Golfkrieges verwickelte die Öffentlichkeit in eine Situation, die alle Definitionsmerkmale der Doppelbindung aufweist: infolge ihres Informationsmonopols bestand eine intensive Abhängigkeit von den Medien, die es besonders wichtig machte, deren Mitteilungen genau zu verstehen, um angemessen darauf reagieren zu können. Mit dem Anspruch, realitätshaltige Informationen zu liefern, bei gleichzeitiger Mitteilung, daß diese Informationen infolge der Zensur unglaubwürdig seien, übermittelten die Medien jedoch zwei widersprüchliche Botschaften, zu denen man wegen des Mangels an unabhängigen Informationsquellen nicht Stellung beziehen, sich aber infolge der Allgegenwart des Krieges auch nicht aus der Situation zurückziehen konnte ..."

Kempf & Reimann 1994, S. 47

# 1 "Ein bißchen Frieden" - Krieg und Frieden in der Populärkultur

# 1.1 Zur Entstehung des Themas

Das Thema Krieg und Frieden ist vor allem in den Nachrichtenmedien präsent; die mediale Beschäftigung mit Konflikten, Krisen und Kriegen findet primär in den Nachrichtenteilen der Printmedien, in den Nachrichtensendungen von Funk und Fernsehen und auf entsprechenden Seiten von Internet und Online-Diensten statt. Wilhelm Kempf und die Projektgruppe Friedensforschung Konstanz haben mit dem Modell eskalations- vs. deeskalationsorientierter Berichterstattung (Kempf et al. 1996) und der nachfolgenden Forschung viel dazu beigetragen, die Mechanismen zu fassen und zu beschreiben, mittels derer die scheinbar an objektiven Kriterien orientierte Konfliktberichterstattung der Nachrichtenmedien zur Eskalation der Konflikte beitragen und zur einseitigen Kriegsberichterstattung werden kann.

Für uns, die wir mit Wilhelm Kempf zu diesen Themen zusammen gearbeitet haben, beide aber auch als (Hobby-)Musiker aktiv und an populärer Kultur und Musik stark interessiert sind, stellte sich die Frage, wie mit dem Thema Krieg und Frieden im Pop und im Schlager umgegangen wird. Zwar sind es die Nachrichtenmedien,

die die öffentliche Meinung zu konkreten Konflikten bilden und formen – Nachrichtenmedien werden in Kriegszeiten vermehrt genutzt und während aktueller Krisen oder Wendepunkte steigt das Volumen der Konfliktberichterstattung zum Teil so stark an, dass man ihr fast nicht mehr entkommen kann (Kempf & Reimann 1994); dennoch nehmen im Alltag der meisten Menschen andere Formate einen viel breiteren Raum ein: Unterhaltungssendungen werden im Allgemeinen von mehr Menschen und länger konsumiert als Nachrichtensendungen, so dass fraglos – zumindest in Friedenszeiten – von einer Allgegenwart der Unterhaltungsindustrie gesprochen werden kann. Es ist somit anzunehmen, dass die Art und Weise der Darstellung von Krieg und Frieden in der vordergründig so apolitischen Populärkultur für das Verständnis von Konflikten generell und für die Einstellung zu konkreten Kriegen ebenso prägend ist wie die konkrete Berichterstattung der Nachrichtenmedien darüber.

Der Eurovision Song Contest (ESC) drängt sich für eine solche Untersuchung geradezu auf. Zum einen ist er eines der größten Medienereignisse der europäischen Populärkultur überhaupt. Zum anderen werden hier inhaltlich zwar vorrangig Themen wie Herzschmerz, Liebe, Jugend und Musik behandelt; dennoch kommen auch die Themen Frieden und, in geringerem Ausmaß, Krieg bzw. Konflikt vor und werden so von einem Millionenpublikum rezipiert. Das Thema Frieden kann sogar großen Erfolg haben: Der Gewinnertitel von 1982 war "Ein bißchen Frieden", der mittlerweile einer der bekanntesten deutschen Schlager überhaupt ist.

Im Folgenden stellen wir unsere Auseinandersetzung mit der Behandlung des Themas Krieg und Frieden im ESC vor. Ausgehend von der Annahme, dass der ESC von den teilnehmenden Ländern als Plattform zur nationalen kulturellen Selbstdarstellung im Allgemeinen, im Besonderen aber auch zu ihrer Selbstdarstellung als "friedliche" oder "friedensliebende" Nationen, als "Friedensmächte" oder Ähnliches genutzt wird, konnten wir ausreichendes Daten- und Anschauungsmaterial erheben. Dieses Datenmaterial entzog sich zunächst jedoch ein Stück weit unseren weiteren Forschungsintentionen bzw. den von uns verwendeten Kriterien und Variablen des bereits erwähnten Modells eskalations- vs. deeskalationsorientierter Berichterstattung (Kempf et al. 1996). Im Weiteren werden wir daher die Überlegungen darstellen, die uns schließlich zu einer Erweiterung des Modells von Kempf et al. führten: Wir präsentieren eine um viele Textbeispiele angereicherte Checkliste Friedenskitsch, die die Analyse populärer Songtexte über Krieg und Frieden ermöglicht. Zudem erlaubt sie die systematische Beschreibung der in der populären Kultur verwendeten Mechanismen zur Abwehr einer angemessenen Beschäftigung mit diesem Thema.

# 1.2 Der Eurovision Song Contest – Festival der Lieder und Plattform der kulturellen Selbstdarstellung

Alljährlich feiert die europäische Fernsehgemeinde ein Fest internationaler Eintracht und gesungener Harmonie: den Eurovision Song Contest. Der seit 1956 stattfindende Musikcontest – ursprünglich ein Komponistenwettbewerb – ist eine

Gemeinschaftssendung der in der European Broadcasting Union (EBU) zusammengeschlossenen staatlichen und öffentlich-rechtlichen Fernsehsender. Er wurde mit der Intention ins Leben gerufen, einen regelmäßigen Anlass für die Fernsehzusammenarbeit der EBU-Mitgliedsländer zu schaffen. Für die EBU hat er somit eine konstituierende und integrative Funktion (Le Guern 2002).

Seinen Teilnehmern bietet der ESC eine internationale Plattform zur kulturellen Selbstdarstellung. "Besonders für kleine und relativ unbekannte Länder, wie beispielsweise Malta oder die baltischen Staaten, ist die Veranstaltung mit ihrer weiten öffentlichen Beachtung in Europa von immenser Bedeutung. Dasselbe gilt auch für zukünftige EU-Mitglieder und Mitgliedsanwärter" (Schweiger & Brosius 2003, S. 274).

In der Tat nimmt der ESC im alltäglichen Fernsehgeschehen eine singuläre Stellung ein – vergleichbar etwa mit sportlichen Großereignissen –, was sich auch in den international relativ hohen Einschaltguoten niederschlägt (Wolther 2006). Bereits Monate im Voraus füllt das anstehende Ereignis die Seiten der Gazetten, lässt Internet-Foren anschwellen und ist in Fernsehmagazinen und -journalen präsent. Dabei zeichnen den ESC jene Eigenschaften aus, die generell auf Medienereignisse (Dayan & Katz 1992) zutreffen: seine Gestaltung, Erzählung und Wahrnehmung in Form von contest, conquest und coronation (Wettkampf, Eroberung, Siegerehrung) und seine Inszenierung als Ritual. Zu einem Ritual wird er vor allem aufgrund der periodischen Ausstrahlung, seiner seit 50 Jahren annähernd unveränderten Dramaturgie mit festem Ablauf und Regelwerk, seines Vergangenheitsbezugs, der mit jeder Neuauflage aktualisiert wird, und durch den Appell an das Zugehörigkeitsgefühl der Zuschauer – einerseits zur großen Masse einer internationalen Fernsehgemeinde, andererseits zu einem nationalen Kollektiv (Wolther 2006). Durch seine feierliche Inszenierung und die gefühlte (bzw. den regelmäßigen Rekurs auf die) Anwesenheit eines Millionenpublikums wird der ESC, wie Medienereignisse ganz allgemein, zu einem grenzübergreifenden, Gemeinschaft und gemeinsame Identität stiftenden Erlebnis. "Nicht selten bieten diese Übertragungen [von Medienereignissen] die Gelegenheit zur Evokation eines idealisierten Selbstbildes der Gesellschaft, die sich oft über eine Aktualisierung des historischen Gedächtnisses vollzieht. Media Events der genannten Art sind daher für die Autoren [Dayan und Katz] Feiertage, an denen sich Gesellschaften weniger daran erinnern, was sie sind, als vielmehr daran, was sie sein wollen." (Keppler 2000, S. 150)

Für seine Zuschauer wiederum ist der ESC, wie dies Raykoff (2007, S. 1) auf den Punkt bringt, eine "annual music competition Europeans love to hate". Der ESC bewegt eine ganze Bandbreite von Emotionen – von nationalem Stolz (wie 2010 anlässlich des Sieges "unserer Lena", Bücken 2010) bis zu nationaler Scham (z. B. Finnlands Trauma anhaltenden Scheiterns seiner Beiträge, bis der Bann 2006 gebrochen werden konnte; Pajala 2007). Sie reichen von Empörung über Voting-

<sup>1</sup> Man denke z. B. an das Ritual der zweisprachigen Punktevergabe: Werden die Punkte in englischer Sprache verkündet ("Germany – twelve points"), folgt nach jeder Nennung unmittelbar die französische Übersetzung ("Allemagne – douze points"), und umgekehrt.

Blocks (Gatherer 2006) auf der einen Seite bis zum Jubel über empfangene Nachbarschaftshilfe bei der Punktevergabe. Tiefe Seelenrührung (z. B. angesichts des "bubenhaften" Charmes Alexander Rybaks, Sieger 2009; Bönisch & Hauck 2009) steht Seite an Seite mit lustvollem Spaß am Spiel mit neu erfundenen, transnationalen Identitäten und postmodernem Trash und Camp (Raykoff 2007), während andere sich angewidert vom Global Kitsch (LeGuern 2002) distanzieren, ihn bewusst ignorieren oder die Veranstaltung als Ganzes einer wissenschaftlichen Vivisektion unterziehen.

# 2 Der Forschungsprozess

### 2.1 Forschungsleitende Fragen

Insbesondere der Aspekt der Evokation eines idealisierten Selbstbildes (Dayan & Katz 1992) war es, der bei uns die Frage aufwarf, wie sich solcherart umfassende Themen und grundlegende kollektive Erfahrungen wie Frieden und Krieg in den Beiträgen des ESC niederschlagen. Prototypisch hierfür steht das bereits erwähnte Lied "Ein bißchen Frieden" der deutschen Sängerin Nicole. Ihr Sieg im Jahr 1982 war für die Bundesrepublik Deutschland nicht etwa nur deshalb ein großer, fast historischer Erfolg, weil alle bisherigen Versuche, den Contest zu gewinnen, trotz guter Platzierung gescheitert waren oder weil hier ein besonders gutes Lied oder eine gute Performance geglückt wären. Vielmehr war es gelungen, zur rechten Zeit das rechte Thema zu präsentieren und damit national wie international einen (gemeinsamen) Nerv zu treffen. Deutschland konnte sich sowohl mit diesem Lied als auch mit der Sängerin und ihrer Selbstdarstellung als "Friedensengel" nach innen wie nach außen als harmlos, friedfertig und friedensliebend darstellen. Der Sieg schien nur noch zu bestätigen, dass diese Botschaft im internationalen Umfeld auch so angekommen war<sup>2</sup>.

Den Beiträgen sind gemäß den Regeln des ESC Texte, Ansprachen und Gesten politischer Natur zwar explizit untersagt (ebenso Texte oder eine Bühnenshow, die den Wettbewerb allgemein in Misskredit bringen könnten, vgl. *Rules for the Eurovision Song Contest 2010*, http://www.eurovision.tv/upload/esc2010rules.pdf); eine erste Recherche ergab aber, dass Frieden und Krieg regelmäßig in den Wettbewerbsbeiträgen thematisiert werden. Uns interessierte, wie diese Themen in den Texten verarbeitet werden. Wie lässt sich zum Beispiel über Frieden und Krieg singen, wenn deren politische Dimension ausdrücklich ausgeklammert wird? In welchen Rahmen werden diese Themen eingebettet? Aus welcher Perspektive werden sie besungen? Werden Ursachen oder Auswirkungen benannt? Welche Lösungen werden angeboten?

Für eine vergleichbare Politur des deutschen Images sorgten erst wieder die Austragung der Fußballweltmeisterschaft 2006 in Deutschland unter dem Motto "Die Welt zu Gast bei Freunden" sowie der Achtungserfolg der deutschen Fußballnationalmannschaft bei der WM in Südafrika 2010: Stellvertretend für das ganze Land wurde die Mannschaft als jung, dynamisch und bunt wahrgenommen. Dieses "Verdienst" wurde denn auch symbolisch mit der Verleihung des Bundesverdienstkreuzes an den Trainer honoriert.

#### 2.2 Textauswahl

Die Textauswahl erfolgte durch eine Volltextsuche mithilfe der Suchmaschine Google auf der Website *The Diggiloo Thrush* (http://www.diggiloo.net). *The Diggiloo Thrush* ist eine nichtkommerzielle Website, auf der die Texte aller Beiträge, die jemals beim Eurovision Song Contest eingereicht wurden, zusammen mit ihrer englischen Übersetzung verfügbar sind. Unsere zentralen Suchbegriffe waren Frieden und Krieg sowie die englischen Pendants peace und war.

Die Suche umfasste den Zeitraum von 1956 bis 2009. In die Untersuchung gingen, unabhängig davon, ob sie den Weg ins Finale schafften, alle für den Wettbewerb eingereichten Texte ein, ausgenommen diejenigen, die vom Land wieder zurückgezogen worden waren, oder Alternativtexte, die im Zuge der weiteren Vermarktung des Stückes entstanden. Die Vollständigkeit und Korrektheit der Suchergebnisse wurde per Handsuche nachgeprüft. Dabei wurden Texte ausgesondert, in denen die Suchbegriffe nur als Redewendung oder im metaphorischen Sinne verwendet wurden, z. B. "war" als Bild für schwierige Liebesbeziehungen, oder "peace" im Sinne von Stille oder persönlichem Seelenfrieden.

Jahr	Land	Songtitel original (englische Übersetzung)	Keyword Frieden/ peace	Keyword Krieg/war	Haupt- thema
1974	Jugo- slawien	Generacija '42 – Moja generacija (Generation '42 – My generation)		x	
1974	Norwe- gen	Hvor er du? (Where are you)	х		х
1976	Luxem- burg	Chansons pour ceux qui s'aiment (Songs for those who love each other)	х		
1977	Frank- reich	L'oiseau et l'enfant (The child and the bird)		х	
1979	Öster- reich	Heute in Jerusalem		х	х
1979	Spanien	Su canción (Your song)	х		
1979	Finnland	Katson sineen taivaan (I'm looking at the blue sky)	х		
1980	Marokko	Bitaqat hob (Love card)	х	х	х
1981	Frank- reich	Humanahum		х	
1982	Deutsch- land	Ein bißchen Frieden	х		х
1985	Israel	Ole, ole (Going up and down)	х		
1986	Türkei	Halley	х		
1987	Belgien	Soldiers of love	х	х	х

Jahr	Land	Songtitel original (englische Übersetzung)	Keyword Frieden/ peace	Keyword Krieg/war	Haupt- thema
1989	Öster- reich	Nur ein Lied		х	х
1990	Portugal	Há sempre alguém (There's always someone)	x		
1993	Kroatien	Don't ever cry	х		х
1993	Nieder- lande	Vrede (Peace)	x		х
1993	Italien	Sole d'Europa (Sun of Europe)		х	х
1994	Öster- reich	Für den Frieden der Welt	х		х
1995	Kroatien	Vukovi umiru sami (Wolves die alone)		х	х
1995	Russland	Kolibjeljnaja dlja vulkana (Lullaby for a volcano)	х	х	х
1996	Israel	Shalom Olam (Peace world/Hello world)	х		х
1996	Irland	The voice	х		
1997	Norwe- gen	San Francisco	х	х	
1999	Deutsch- land	Reise nach Jerusalem – Kudüs'e seyahat (Journey to Jerusalem)	х		
1999	Frank- reich	Je veux donner ma voix (I want to give my voice)		х	
2002	Frank- reich	Il faut du temps – Je me battrai pour ça (It takes time – I will fight for that)	х	х	х
2003	Polen	Keine Grenzen-Żadnych granic (No borders)	х	х	х
2004	Frank- reich	À chaque pas (With every step)	х		
2005	Lettland	The war is not over		х	х
2005	Finnland	Why	х		
2006	Zypern	Why angels cry	х		
2007	Israel	Push the button	х	х	х
2008	Geor- gien	Peace will come	х	х	х
2008	Schweiz	Era stupendo (It was wonderful)	х		
2009	Nieder- lande	Shine		x	х

Tabelle 1: Textauswahl mit Schlüsselbegriffen

Die Suche nach den Schlüsselbegriffen ergab 69 Texte. Davon enthielten 46 das Wort Frieden/peace, 14 das Wort Krieg/war, 9 enthielten beide. Ausgeschlossen wurden 33 Texte, die die Suchbegriffe als Metaphern für einen anderen Sachverhalt oder als Redewendung enthielten. Von den letztlich ausgewählten Texten (vgl. Tab. 1) beschäftigen sich 19 unmittelbar mit Frieden und Krieg bzw. der Möglichkeit, Frieden zu schaffen; in den übrigen Songs sind diese Themen Nebenschauplätze. Es fällt auf, dass Frieden und Krieg erst knapp 20 Jahre nach der ersten ESC-Ausstrahlung besungen werden, danach jedoch einen festen Platz einnehmen. Ebenso auffällig ist die Häufung von Texten, die sich zentral mit dem Thema beschäftigen, von Anfang bis Mitte der 1990er-Jahre. Dies lässt sich in unmittelbarem Zusammenhang mit den Kriegen im ehemaligen Jugoslawien sehen.

#### 2.3 Methoden der Zusammenarbeit

Da es sich beim Forschungsteam Jaeger/Reimann um eine für das Verfassen dieses Beitrags ad hoc gebildete Arbeitsgruppe handelt, mussten wir Mittel und Wege finden, um trotz räumlicher Entfernung und hoher zeitlicher Belastung durch Beruf und Familie effizient zusammenarbeiten zu können. Wir richteten daher im Internet ein eigenes, nur uns zugängliches Wiki mit dem Arbeitstitel *Peace and Conflict in the European Song Contest* auf einem privat gemieteten Server ein. Ein Wiki ist "ein Hypertext-System für Webseiten, dessen Inhalte von den Benutzern nicht nur gelesen, sondern auch online direkt im Browser geändert werden können" (Wikipedia, Artikel Wiki). Diese Eigenschaft machten wir uns zunutze und verwendeten das Wiki zunächst zur Datensammlung, später für Kodierungen und Datenanalyse sowie zur Entwicklung der Ergebnisse und zur (vorläufigen) Formulierung dieses Beitrags.

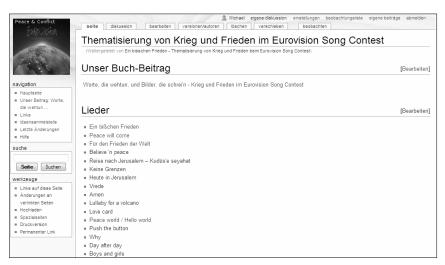


Abbildung 1: Wiki Peace and Conflict in the European Song Contest

Ein Wiki bietet gegenüber dem gemeinsamen Schreiben an einer offline-Datei, wie zum Beispiel einer Word-Datei, mehrere wertvolle Vorteile, beispielsweise

- Kommentarfunktion: Alles Geschriebene kann vom anderen kommentiert werden.
- Neue Seiten: Es können jederzeit neue Seiten beispielsweise zum Sammeln oder Entwickeln neuer Ideen eingerichtet werden.
- Verlinkungen: Es können interne wie externe Links auf weitere Texte, Literatur oder auch, wie in unserem Falle, Videos gesetzt werden.
- Versionierung: Alle Versionen der Seiten werden dauerhaft gespeichert, so dass jederzeit zu einer früheren Version zurückgegangen werden kann.

# 2.4 Analysemethoden

Aufgrund unserer Erfahrungen bei der Inhaltsanalyse von Konfliktberichterstattung schien es uns zunächst naheliegend, als Untersuchungsinstrument für Friedens- oder Kriegstexte auf das Modell eskalations- vs. deeskalationsorientierter Berichterstattung (Kempf et al. 1996) zurückzugreifen und uns an dessen Analyseebenen zu orientieren. Ausgangspunkt der von uns geplanten quantitativen und qualitativen Inhaltsanalysen war die Verwendung der Variablen der auf dem Modell basierenden *Checklist for escalation- and de-escalation-oriented aspects of conflict coverage* von Bläsi et al. (2004). Diese hatte sich in ihrer erweiterten Version nicht nur für die Analyse von Konfliktberichterstattung, sondern auch für die Evaluation von Zeitungstexten in Nachkriegs- und in Friedenssituationen bewährt (z. B. Jaeger 2009).

Natürlich war aufgrund der Thematik und des Rahmens eines festlichen Medienereignisses kaum zu erwarten, Textbeispiele auf dem eskalationsorientierten Pol der Checkliste zu finden. Zudem handelt es sich beim Untersuchungsmaterial nicht um Medienberichterstattung, sondern um Songtexte bzw. Lyrik. Es schien uns jedoch einen Versuch wert, zu überprüfen, inwieweit sich das Analyseschema auch für die Untersuchung alternativer Textsorten eignet und wo es an Grenzen stößt. Vor allem die logische Struktur der verschiedenen Konfliktebenen (von der Konzeptualisierung der Situation bis hin zu Anreizen zu sozialer Identifikation) war es, die das Unterfangen realisierbar erscheinen ließ.

# 2.5 Neuorientierung

Es stellte sich jedoch bald heraus, dass die für Nachrichtenberichterstattung formulierte und dort auch bewährte Checkliste auf die Songtexte des ESC kaum anwendbar ist. Zum einen fehlt in den meisten Texten der Bezug auf einen konkreten Konflikt. Damit entfiel die Grundlage für die meisten aus dem Modell abgeleiteten Variablen; z. B. ließen sich in unserer Stichprobe so gut wie nie konkrete Konfliktparteien identifizieren, so dass die Codierung sämtlicher Variablen mit Bezug auf die eine oder andere oder beide Konfliktparteien von vornherein nicht möglich war. Zum anderen fiel auf, dass die Songtexte ein Element enthalten, das wir zunächst

nur vage greifen und beschreiben konnten. Es war das Element der Idealisierung, Schönfärberei, Verallgemeinerung und Übertreibung, für das wir letztlich den Begriff Kitsch fanden.

Literaturrecherchen zum Thema Kitsch bestätigten unsere Vermutung. Milan Kundera beispielsweise schreibt in seinem Roman *Die unerträgliche Leichtigkeit des Seins*: "Der Kitsch ruft zwei nebeneinander fließende Tränen der Rührung hervor. Die erste Träne besagt: wie schön sind doch auf dem Rasen rennende Kinder! Die zweite Träne besagt: wie schön ist es doch, gemeinsam mit der ganzen Menschheit beim Anblick von auf dem Rasen rennenden Kindern gerührt zu sein! Erst diese zweite Träne macht den Kitsch zum Kitsch. Die Verbrüderung aller Menschen dieser Welt wird nur durch den Kitsch zu begründen sein." (1987, S. 240)

Kundera bringt damit die wichtigsten Merkmale von Kitsch auf den Punkt: Die Essenz des Kitsches ist es, Emotionen zu erzeugen bzw. es Kitsch-Nutzern zu ermöglichen, sich an den eigenen Emotionen zu erfreuen und sich selbst zu feiern. Nach Kulka (1996) werden im Kitsch Objekte oder Themen dargestellt, die mit kulturell geteilten Erfahrungen und Emotionen hoch aufgeladen sind. Diese Objekte oder Themen sind anhand einfacher Schlüsselreize unmittelbar zu identifizieren und erlauben mühelosen Genuss. Kitsch trägt insofern nichts dazu bei, unsere bislang mit den Objekten oder Themen verbundenen Assoziationen substanziell zu bereichern. Im Kitsch wird Atmosphäre verkauft; es geht nicht um die ästhetische Rezeption des Objekts, sondern um die der dadurch angesprochenen Gefühle. Kitsch bedient sich in parasitärer Weise der damit verbundenen, leicht zugänglichen Assoziationen, um emotionale Reaktionen zu erzeugen. Dabei gibt er sich ernsthaft und ist völlig frei von Ironie<sup>3</sup>. In der Konseguenz verstärkt Kitsch kulturell eingebettete Überzeugungen, wie die Welt organisiert ist und welchen Platz man in ihr einnimmt. Kitsch verzichtet auf die Analyse der realen Verhältnisse, er baut sie stattdessen neu zusammen und stilisiert sie (Higgins 2009).

#### 2.6 Friedenskitsch

Da diese Charakterisierungen auf einen Großteil unseres Datenmaterials zutreffen, kamen wir zu dem Schluss, dass die Checkliste von Bläsi et al. für die Anwendung auf "kitschiges" Datenmaterial zu modifizieren sei. Es wurde klar, dass unseren Songtexten nicht mit der an einer (oberflächlich) sachlichen Medienberichterstatung orientierten Checkliste beizukommen ist, weil es sich bei den "friedensbewegten" Texten, die im Rahmen des ESC vorgetragen werden, überwiegend um Friedenskitsch handelt. Zu Kriegskitsch, einer Kategorie, nach der wir ebenso Ausschau hielten, fanden sich leider so gut wie keine Texte.<sup>4</sup>

<sup>3</sup> Im Unterschied dazu werden unter Camp-Style Darbietungen oder Objekte verstanden, deren Reiz darin liegt, dass Kitsch absichtlich und mit Ironie eingesetzt wird.

<sup>4</sup> Als Ausnahme lässt sich "Dschinghis Khan" nennen. In diesem Text wird das Leben und Wirken des gefürchteten Kriegsherren romantisch verklärt.

Zwar wird das Thema Frieden, oder der Wunsch danach – und implizit damit auch das Thema Krieg –, beim ESC in "getragener" Weise behandelt oder mit dem Anspruch darauf, zum Nachdenken anzuregen; auf der anderen Seite stellen die Texte durchweg verschiedenste Spielarten von Friedenskitsch zur Schau, die eine authentische Nachdenklichkeit oder gar dem Thema angemessene Reaktionen sofort wieder abwehren bzw. verhindern. Friedenskitsch vermittelt beispielsweise, dass gegen Kriege nichts auszurichten sei, dass Frieden unerreichbar sei; dass man gegen Kriege und die Kriegführenden eigentlich nur ansingen könne, am besten gemeinsam; er vereinfacht Konflikte als Kampf zwischen Gut und Böse und feiert zwar die "Friedensliebenden", die jedoch nichts wirklich erreichen. Diese und weitere Spielarten von Friedenskitsch stellen wir in Kapitel 3 detailliert und anhand von Beispielen vor.

### 2.7 Konstruktion der Checkliste Friedenskitsch

Friedenskitsch definieren wir über die Formulierungen der Variablen auf den Analyseebenen Konzeptualisierung der (Konflikt-)Situation, Rechte und Intentionen, Handlungen, Emotionale Verwicklung sowie Anreize zu sozialer Identifikation des kempfschen Modells bzw. der bläsischen Checkliste (s. u.). Die Analyseebenen wurden also beibehalten. Die Ausformulierung der einzelnen Variablen musste jedoch so angepasst werden, dass auf jeder Ebene das dem Datenmaterial eigene Element des Kitsches definiert und damit greif- und operationalisierbar wurde. Aus der Auseinandersetzung mit der Literatur ergab sich notwendigerweise, dass dieses Kitsch-Element durch Stilmittel wie Übertreibung, Überhöhung, Idealisierung, Verherrlichung oder Romantisierung, stellenweise auch Verniedlichung und Verharmlosung zu definieren war.

Die ursprüngliche Checkliste ist zweigeteilt, sie stellt eskalationsorientierte Aspekte von Konfliktberichterstattung den deeskalationsorientierten gegenüber. Auf der Grundlage der deeskalationsorientierten Aspekte konnten nun mittels Übertreibung, Idealisierung und Schönfärberei diejenigen Variablen formuliert werden, die wir als Friedenskitsch bezeichnen. Obwohl auf den ersten Blick absurd erscheinend, ergab auch die Einbeziehung der eskalationsorientierten Aspekte einige zusätzliche Anregungen zur Formulierung von Kitschvariablen. Insgesamt stellte sich aber heraus, dass eine Überführung der ursprünglichen Variablen durch Übertreibung oder Verniedlichung zusätzliche wesentliche Aspekte von Kitsch unberücksichtigt ließ, z. B. die Stimulierung eines Wohlgefühls bei den Rezipienten. Wir mussten die Variablen daher zum Teil neu formulieren, indem wir sowohl Beobachtungen am Datenmaterial als auch theoretische Überlegungen zum Element des Kitsches berücksichtigten.

<sup>5</sup> Man denke an Nicole und ihre Selbstinszenierung als das einsame, nachdenkliche Mädchen mit der Gitarre

Es gibt auch die Variante, dass der Text sehr wohl ernsthafte und angemessene Aussagen trifft, die Darbietung (Performance) und die Musik jedoch so gefällig und lieblich sind, dass die Schärfe des Textes konterkariert und eine mögliche Wirkung wieder aufgehoben wird (z. B. "The war is not over", gesungen von Walters & Kazha, Lettland, 2005).

In einem nächsten Schritt – und danach alternierend immer wieder – legten wir die so entwickelte *Checkliste Friedenskitsch* an das Datenmaterial an und suchten Beispiele für die theoretisch erarbeiteten Variablen. Die Formulierungen der Variablen konnten so überprüft und geschärft werden. Zum Teil verlief der Prozess auch in der anderen Richtung: Aus dem Datenmaterial ergaben sich Anhaltspunkte und Ideen für weitere Teilaspekte von Variablen, die wiederum in verallgemeinerter und abstrahierter Form in die Checkliste eingingen. Das Forschungsprozedere (nicht nur, aber besonders) in dieser Phase lehnte sich damit an das Forschungsparadigma der Grounded Theory an (Glaser & Strauss 1967); zum einen durch die Anwendung des theoretischen Samplings, um zielgerichtet weitere Textbeispiele zu finden, zum anderen durch den Wechselschritt der Methode des permanenten Vergleichs zwischen Daten und Theoriebildung. Ergebnis dieser Arbeit ist die *Checkliste Friedenskitsch* (vgl. Kap. 3.2).

# 3 Ergebnisse

# 3.1 Analyseebenen von Friedenskitsch

Konzeptualisierung der Situation: Im ursprünglichen Analyseschema betrifft dies die Frage, wie die (Konflikt-)Situation konstruiert wird – als Konkurrenzprozess oder als kooperativer Prozess. Unsere Songtexte über Frieden oder Krieg hingegen enthalten so gut wie nie konkrete Konfliktsituationen. Vielmehr geht es oft um den unauflöslichen Widerspruch der Sehnsucht nach Frieden in einer unfriedlichen Welt. Dieser wird zugespitzt auf einen "Kampf" der verherrlichten "Guten", die sich nach Frieden sehnen und dafür einsetzen, gegen das dämonische, allgegenwärtige, aber dadurch auch wenig greifbare "Böse" in Form von Unrecht, Leid, Gewalt, aber auch nur gegen schlechte Gefühle angesichts des Zustands der Welt. Nur eine Seite kann siegen. Doch auch das Gegenteil, die Dekonstruktion des unauflöslichen Antagonismus, gerät in die Nähe von Friedenskitsch, wenn nämlich der gesamten Menschheit romantisierend unterstellt wird, nichts weiter zu wollen als in Frieden und Harmonie zu leben, während Krieg und Gewalt scheinbar vom Himmel fallen.

Friedenskitsch appelliert an kollektive Sehnsüchte nach einer heilen, vollkommenen Welt, die in der Vergangenheit angeblich bestand, in der Gegenwart bzw. im aktuellen Moment gelebt wird, in der Zukunft zu gestalten oder als unerreichbarer Traum nur durch die Intervention himmlischer, göttlicher Gewalt zu erreichen ist. Frieden, Zustände des Friedens und die Möglichkeiten, Frieden zu schaffen, werden verherrlicht und in romantisch-schöne Farben getaucht; teilweise aber auch verniedlicht und auf einfache Formeln gebracht, z. B. "lieben statt hassen". In gleicher Weise werden Kriege, Konflikte und Gewalt überzeichnet: Sie werden von ihren Folgen her beschrieben, erzeugen Leid und bei denen, die den Frieden lieben, ungute Gefühle einer kollektiven, vagen Bedrohung und einer grundlegenden Trauer über die schlechte Welt. Teilweise sind sie auch ein Ausdruck des Bösen schlechthin und dem Einfluss höherer, finsterer Mächte (des Bösen allgemein) oder egoistischer Menschen zuzuschreiben, die nie konkret benannt werden.

Rechte und Intentionen: Diese Variablen befassen sich generell damit, was jemand zu erreichen versucht, mit übergeordneten Zielen und Rechten auf einer eher allgemeinen, abstrakten Ebene (Kempf et al 1996). Bei den eskalationsorientierten Elementen der ursprünglichen Checkliste stehen sich idealisierte eigene Ziele und Rechte und dämonisierte Absichten und die Verweigerung von Rechten auf der anderen Seite gegenüber. Deeskalationsorientierung hingegen entspricht einer unvoreingenommenen Darstellung der Ziele und Rechte der anderen bzw. einer realistischen, selbstkritischen Darstellung eigener Absichten und Rechte.

Friedenskitsch findet sich auf dieser Ebene in der Idealisierung der eigenen Friedfertigkeit, Friedensabsichten und -sehnsüchte. Die Befürworter des Friedens ("us") vertreten die Seite des Guten. Wer Krieg führt ("them"), zählt hingegen zum Bösen und wird entsprechend dämonisiert. Im Grunde entspricht dies dem eskalationsorientierten Pol der ursprünglichen Checkliste (eigene Idealisierung vs. Dämonisierung der anderen), so dass sich hier auch von Friedenspropaganda sprechen ließe. Kitschig daran ist die märchenhafte Verkürzung auf den antagonistischen Kampf zwischen Gut und Böse, wobei es dem Zuhörer einfach gemacht wird, sich auf die Seite der Guten zu schlagen. Die "Bösen" werden oft nur in Andeutungen beschrieben. Sie bleiben namen- und gesichtslos, ihre Intentionen wenig konkret greifbar. Das vorgebliche Ziel, Frieden erreichen zu wollen, muss sich somit darauf beschränken, sich in dem Bewusstsein der eigenen moralischen Größe und im Verbund mit Gleichgesinnten wohl zu fühlen, ohne dass dies – abgesehen von Liedersingen und einander lieb haben – konkrete Handlungen zur Zielerreichung nach sich ziehen müsste. Die "Kriegstreiber" bleiben unerreichbar, weil unerkannt. Was letztlich zählt, ist die globale Verbundenheit im ozeanischen Gefühl, auf der richtigen Seite zu stehen, wozu auch die Betonung und Verklärung weltumspannender gemeinsamer Friedenssehnsucht beiträgt. Kritisches Hinterfragen der eigenen Ziele und Rechte beschränkt sich seitens der Ich- oder Wir-Erzähler in den Texten auf vordergründige Selbstkritik, die die eigene moralische Größe nur zusätzlich untermauert.

Da Songtexte oft auch Appelle an die Zuhörer beinhalten, zählen wir zu Friedenskitsch auf der Ebene der Intentionen außerdem solche Aufrufe, in denen das Publikum aufgefordert wird, sich mehr um den Frieden und das Gute zu bemühen (ohne konkreten Anhalt, welche Handlungen hierfür notwendig wären).

Handlungen: Wie werden die Mittel beschrieben, die zur Erreichung der Ziele eingesetzt werden? Wiederum geht es bei Friedenskitsch primär um romantisierende oder verklärende Selbstdarstellungen im (vorgeblichen) Einsatz für den Frieden. So wird das eigene Potenzial, Frieden zu schaffen, übertrieben und überschätzt. Im Prinzip triviale Handlungen (gemeinsames Singen und Feiern) werden zu Aktivitäten stilisiert, die dem Frieden dienen. Man feiert sich selbst. Hierzu zählt auch die begeisterte Beschreibung gemeinsamen "Friedenshandelns" im jetzigen Augenblick, überall auf der Welt, indem man zusammen singt und sich miteinander wohlfühlt. Zu Friedenskitsch zählen wir außerdem vorgebliche Selbstkritik im Sinne einer Überhöhung der eigenen Harm- und Hilflosigkeit: Es wird festgestellt, dass sich die eigenen Handlungsoptionen aufs Singen, Musizieren und Appellieren be-

schränken – was man dafür mit umso mehr Herzblut betreiben muss. Wenn es auch gegenüber der übermächtigen Bedrohung nichts nützen mag – relevant ist die richtige Attitüde. Hierzu gehört es auch, die Zuhörer zum Mitmachen aufzufordern (in der aktuellen Situation des Liedvortrags wie auch generalisierend), nämlich z. B. für den Frieden zu singen; seinen inneren Frieden zu finden, um ihn in die Welt zu tragen; gegen Unrecht anzuschreien; hier und heute ein Friedensfest zu feiern etc.

Emotionale Verwicklung. Friedenskitsch hat unmittelbar die Erzeugung von emotionalen Antworten bei den Rezipienten im Visier. Es geht primär um eine Steigerung des Wohlbefindens (Behaglichkeit, eigene Korrektheit, Genuss der eigenen Empfindsamkeit, Wir-Gefühl) und der Solidarität mit dem/der Sympathie für den Sänger (dessen Figur/der z. B. an der Welt leidet oder für das Gute kämpft) bzw. mit der Veranstaltung insgesamt. Im Kontext des Themas Frieden und Krieg treten selbstverständlich auch Opfer auf. Ihre Darstellung bleibt auf bekannte, bewährte Klischees reduziert ("sterbende Kinder", "weinende Frauen"). Der Appell an die Rezipientenbetroffenheit bleibt folglich oberflächlich, etwaiger Handlungsdruck kann abgewehrt werden und wird in Ungefährliches, Harmloses kanalisiert.

So fördert ein Gleichgewicht aus Bedrohlichkeit der Situation und Hervorhebung eigener moralischer und zahlenmäßiger Stärke die Bereitschaft, sich auf die Seite der Guten bzw. des Sängers zu schlagen und emotional mitzugehen. Die Darstellung globalen Unrechts in Verbindung mit der Betonung der eigenen Empfindsamkeit demgegenüber bestätigt die eigene höherwertige Position. Vage Opfer- bzw. Unrechtsdarstellungen durch Klischees und demonstratives Bekunden eigenen Mitgefühls sorgen für eine vordergründige Empörung gegenüber Krieg und Unrecht, ohne dass ein echter Handlungsdruck entsteht. Die Betonung der gemeinsamen Erfahrung im "korrekten" Empfinden, Wollen und Handeln fördert die Erzeugung eines unreflektierten Gemeinschaftsgefühls, das sich in spontanem Wohlgefühl erschöpft. So zielt die emotionale Verwicklung bei Kitschtexten stets darauf ab, ein wenig auf die Tränendrüse zu drücken, zugleich aber ein Empfinden moralischer Erhabenheit und des beseelten Eins-seins mit der Welt hervorzulocken.

Anreize zu sozialer Identifikation: Auf dieser Ebene erfahren die Rezipienten, wie die eigene Gruppe oder Gesellschaft beschaffen ist, wofür sie steht, worin sie sich von anderen unterscheidet und was ihr Streben für die Zukunft ist (Kempf et al 1996). Friedenskitsch in Songtexten gibt den Rezipienten primär Anreize für eine Identifikation mit der Figur des Ich-Erzählers (welcher gut ist bzw. das Gute will), mit der Figur des Interpreten, der mit dem Ich des Textes zu einer Person verschmilzt, und – was das Wichtigste ist – mit der Gemeinschaft der Rezipienten. Dies geschieht, indem die Gefühlswelt der singenden Figur, ihre Hoffnungen, Sehnsüchte, aber auch Sorgen, ihr Leid und ihre Enttäuschungen in den Vordergrund gerückt werden und direkt oder indirekt an die Ähnlichkeit dieser Gefühlslagen mit denen der Rezipienten appelliert wird. Das Leid des Sängers steht dann stellvertretend für das Leid aller Menschen auf dieser Erde. Die imaginierte Gemeinschaft aus Hauptfigur, Interpret und Millionen-Publikum steht auf der Seite des Guten und verurteilt das Böse. Um diese fiktive Solidarisierung zu verstärken, wird in einigen Texten zusätzlich auf bekannte Friedensvertreter referiert (z. B. Ghandi). Etwas

vom Glanz ihrer historischen Größe färbt damit auch auf das Publikum ab.

Anreize zur Identifikation mit den Leidtragenden von Krieg und Gewalt beschränken sich hingegen auf bewährte Klischees und vage, blumige Leidensbeschreibungen und bleiben deshalb blass. Zu deutliche Identifikationsanreize mit ihnen könnten den harmonischen Abend verderben. Ähnlich verhält es sich mit der Dehumanisierung von Kriegführenden. Das "Böse" bzw. die "Bösen" werden in Metaphern und vage Umschreibungen gekleidet. Es gibt somit keine echte Adresse für die friedensbewegten Anliegen – außer der eigenen Seite, die sich im Gut-sein feiert. Als Alternative werden die "Bösen" mit derart negativen Attributen dehumanisiert, dass wiederum nur noch eine höhere Gewalt, die man in trauter Gemeinschaft beschwört, Hilfe verspricht.

#### 3.2 Die Checkliste

#### Checkliste Friedenskitsch

#### Konzeptualisierung

Konstruktion als Spannungsfeld Gut gegen Böse, d. h. Frieden gegen Krieg

- Situation als Kampf des/der Guten (für Frieden) gegen das Böse (Krieg/Gewalt/Unrecht etc.)
   It takes time (I will fight for that), Frankreich 2002
   Shine. Niederlande 2009
- Das "Böse"/Krieg/Gewalt/Unrecht etc. als selbstverständlich, allgegenwärtig, das Gute im Kampf dagegen
  - "There are so many wars right now in this world / There are so many things not right / .../ Time is now to make the change / Time is on our side / Love will make us glow in the dark" (Shine, Niederlande 2009)
- Auflösung aller Gegensätze: Alle Menschen sind gut/gleich/friedlich etc. das Böse/ Krieg/Gewalt/Unrecht ist unmenschlich/nicht von Menschen gemacht
   "The child of the earth is looking for love / Trying to hide her longing / All my wishes of the dawning / will be destroyed by the sunset" (I'm looking at the blue sky, Finnland 1979)

#### Mystifizierung / Verherrlichung Frieden

- Frieden als makelloser, paradiesischer Zustand
  - "Hey I see understanding, Hey I see more hope, Hey the darkness disappears. A great light is slowly covering the world" (Peace world / Hello world, Israel 1996)
  - "A dream of peace, a spring of youth / No guns, no war, no disco / Just lovely flowers in your hair" (San Francisco, Norwegen 1997)
  - "Seht, aus der Kühle hoher Luft / da fallen hundert Monde / Als Zeichen für den Neubeginn / Als Sinn wider den Widersinn" (Heute in Jerusalem, Österreich 1979)
- Frieden als jenseitiger, nicht mit menschlicher Macht bzw. nur durch die Intervention höherer Mächte erreichbarer Zustand
  - "A place to pray for peace and harmony" (Why Angels Cry, Zypern 2006)
  - "Wake up, sun, so we can feel you / Today nobody asks for you / Cover Europe with light, do you remember where it is?" (Sun of Europe, Italien 1993)
  - "Hello Halley ... give us some love / Hello Halley ... give us some respect" (Halley, Türkei 1986)
- Frieden als unausweichlicher Zustand
  - "Peace will come" (Peace will come, Georgien 2008)
  - "Nothing can stop love and peace" (There's always someone, Portugal 1990)

#### Simplifizierung von Frieden

- Simplifizierung der Möglichkeit, Frieden zu schaffen bzw. eine Besserung herbeizuführen / Patentrezepte für Frieden und eine bessere Welt
  - "And if we stay friends, a dream will live forever" (Reise nach Jerusalem Kudüs'e seyahat, Deutschland 1999)
  - "Soldiers of love fire off love" (Soldiers of Love, Belgien 1987)
  - "Everyone should go there at least once [ins Weltall] / To realise this struggle doesn't make sense / ... / And instead of talking, we begin to love" (Keine Grenzen-Zadnych granic, Polen 2003)
  - "So the only thing that we have to do / Is love the one we hate" (Shine, Niederlande 2009)
- Verniedlichung / Rosafärbung Frieden
  - "Peace is only a child, that doesn't walk straight" (It takes time (I will fight for that), Frankreich 2002)
  - "Ein bißchen Frieden, ein bißchen Freude..." (Ein bißchen Frieden, Deutschland 1982) "To one another weaving a rainbow, peace has a thousand colours" (Going up and down, Israel 1995)

#### Mystifizierung von Krieg und Gewalt

- Krieg als etwas Übernatürliches, das Wirken einer höheren Macht, eine Naturgewalt
   "Forget forever, volcano, what you have threatened for centuries / Stop to tyrannize the
   world and stay a green mountain" (Lullaby for a volcano, Russland 1995)
- Krieg bzw. Gewalt als übermächtig, unbezwingbar
  - "Why can't love be all that matters / Why? / While we pray to understand / Faith and hope are slipping through our hands" (Why, Finnland 2005)
  - "We build houses to withstand hurricanes / And make ships to sail in every storm / We're working on a light bulb that never burns out / It only just won't work to keep peace" (Peace, Niederlande 1993)

#### Simplifizierung von Krieg und Gewalt

- Umschreibung von Krieg/Gewalt bzw. deren Folgen als Zustand individuellen Unbehagens / als vage Bedrohung / in Form von Metaphern
  - "Worte, die wehtun, und Bilder, die schrei'n" (Für den Frieden in der Welt, Österreich 1994)
  - "Dann seh ich die Wolken, die über uns sind / Und höre die Schreie der Vögel im Wind" (Ein bißchen Frieden, Deutschland 1982)
  - "Yet another house is flattened / A dream buried in dust / A scream rises to the clouds / There is sadness in the sight of the wind" (Where are you? Norwegen 1974)
- Krieg/Gewalt als ur-menschliche Unzulänglichkeit / lässliche Sünde "our stupid megalomania", "no stupid quarrels, no different races" (Keine Grenzen-Zadnych granic, Polen 2003)
- Vereinfachung von Kriegsursachen
   "The men of not long ago were all brothers / But they invented the war" (Humanahum, Frankreich 1981)

#### 2. Rechte und Intentionen

#### Übertreibung und Verherrlichung friedlicher Intentionen

- Idealisierende Darstellung der eigenen Friedfertigkeit und Friedenssehnsucht
   "I wish there was peace and everybody would be hand in hand" (Reise nach Jerusalem Kudüs'e seyahat, Deutschland 1999)
  - "Damit jeder Mensch in Freiheit leben kann / und die Waffen werden Blumen irgendwann / Dafür lohnt es sich, ein Träumer zu sein" (Nur ein Lied, Österreich 1989)
- Überhöhung, Idealisierung der globalen Friedfertigkeit und Friedenssehnsucht / Vereinnahmung aller Menschen (sogar derer, die Krieg führen)
  - "We are the children of the whole world / ... / We didn't want neither wars nor pains / We wanted a life full of peace" (Love card, Marokko 1980)
  - "They change their uniform and colour / But their tired souls wait for the sun" (Sun of Europe, Italien 1993)
- Demonstrative Bescheidenheit hinsichtlich der Ziele
  - "Ein bißchen Frieden, ein bißchen Freude / Ein bißchen Wärme, das wünsch ich mir" (Ein bißchen Frieden, Deutschland 1982)
  - "We are the children of the whole world / asking for a simple happy life" (Love Card, Marokko 1980)
- Beschränkung von Selbstkritik an den eigenen Zielen bzw. an deren Ernsthaftigkeit auf Allgemeinplätze, Unkonkretes oder Altbekanntes
  - "Because there's no peace, as long / As everyone thinks of himself" (It was wonderful, Schweiz 2008)
  - "And here today it could also be this way / We're small, but not necessarily bad (...) Maybe it will take our stupid megalomania" (Keine Grenzen-Zadnych granic, Polen 2003)
- Appelle an die Rezipienten, sich für den Frieden und gegen den Krieg einzusetzen "Say it loud, peace will come / Everybody shout: peace will come" (Peace will come, Georgien 2008)
  - "People, let your light shine / Let the darkness fade away / Step into a brighter day / Don't forget about our world today" (Shine, Niederlande 2009)

#### Dämonisierung / Verdammung

- Verurteilung der Intentionen der "anderen" ("they"), ohne konkrete Adresse
   "It takes time, that's what they say to us, they say / They talk and they write, and the war sells" (It takes time (I will fight for that), Frankreich 2002)
  - "Asking a life free of greed and fire / Free of rival neighbours", "A love card with warm greetings / not sent by the traitors of peace" (Love card, Marokko 1980)
  - "The war is not over, everyone knows it / It's just a reason to make us believe / That someone's the loser, someone's the winner / To make us believe that's the way it should be" (The war is not over, Lettland 2005)
  - "Dark are the misery, the men and the war / Which believe that they hold the time's reins" (The bird and the child, Frankreich 1977)

#### 3. Handlungen

#### Überhöhung eigener "Friedensmacht" und eigenen Friedenshandelns

 Überhöhung eigener Möglichkeiten, Konflikte abzuwenden oder zu lösen (eines Tages werden "wir" Sänger alle Kriege beenden, das Paradies auf Erden herbeiführen etc.)
 "We can redo everything, change the world / To live without frontiers, we'll sing / To forget history, to believe, not to believe anymore / In our memories, it will stay" (Songs for those who love each other, Luxemburg 1976)

- "If there's no more war which is won in advance / If there are no more barriers, no more differences / (...) / If it's enough to sing, if it's enough to sing / I want to give my voice, I want to give my words..." (I want to give my voice, Frankreich 1999)
- Frieden schaffen durch Singen, durch Gemeinsamkeit, durch "Sich wohlfühlen"

  "Maybe the answer is just this simple song / Let everybody shine along / To keep this world from going wrong" (Shine, Niederlande 2009)
  - "So they come to sing tonight from the four corners of the world / The joy in the heart is not enough, a moment of light for all / And meanwhile the song grows louder, bringing the hearts together / Angels spread their wings and the candles are lit" (Going up and down, Israel 1985)
  - "Ich singe nur ein Lied heut' Nacht / Ein Lied, das uns zu Freunden macht" (Nur ein Lied, Österreich 1989)
- Selbstkritik, die die eigene Moral untermauert und nicht schmerzt

   Westernahmen (Norden auf ihren und 16 august 16 au
  - "We feel guilty / but sip our wine / You're asking yourself constantly / What we are doing" (Soldiers of love, Belgien 1987)
  - "While we fight for nothing, my eyes run dry / Are you still so blind to ask me why?" (Peace will come, Georgien 2008)
- Aufrufe zu (trivialen) Handlungen, um Frieden zu schaffen "Singt mit mir ein kleines Lied / Dass die Welt in Frieden lebt" (Ein bißchen Frieden, Deutschland 1982)
  - "Let's sing together without looking backwards / Nothing can stop love and peace" (There's always someone, Portugal 1990)

#### Übertreibung eigener Harm- oder Hilflosigkeit

Beschränkung der eigenen Handlungsmöglichkeiten auf Singen, Musizieren, Appellieren
"Ich weiß, meine Lieder, die ändern nicht viel, Ich bin nur ein Mädchen, das sagt, was
es fühlt ..." (Ein bißchen Frieden, Deutschland 1982)

"I'm only a little soldier with no weapon but my voice" (It takes time (I will fight for that), Frankreich 2002)

#### Verschleierung von Kriegs- und Gewalthandeln

- Beschreibungen von Kriegs- und Gewalthandlungen nur über deren Folgen Ausblenden der Verursacher
  - "Look, the sky is crying cold bitter tears / Weeping for the people lost in fear" (Peace will come, Georgien 2008)
  - "Little light full of ache and yearning / For the child who won't be returning / And all the loved ones left behind" (Why, Finnland 2005)
  - "Children are thrown in the battle / For ideals, for profits" (Soldiers of love, 1987 Belgien)

#### 4. Emotionale Verwicklung

#### Stimulierung von Wohlbefinden

- Wir-Gefühl und Gemeinschaftserleben Feiern der schönen gemeinsamen (aktuellen)
   Erfahrung und der Einigkeit im moralisch korrekten Empfinden, Wollen und Handeln
   "So they come to sing tonight from the four corners of the world / The joy in the heart
   is not enough, a moment of light for all" (Going up and up, Israel 1985)
   "Und heute in Jerusalem / Endlich in Jerusalem / Atem holen und besinnen / Und heute
   in Jerusalem / Friede für Jerusalem" (Heute in Jerusalem, Österreich 1979)
- Nostalgie Beschwörung gemeinsamer (auch historischer) "Friedenserfahrungen" (Mythenbildung)
  - "It still feels like yesterday / When everyone was singing: 'Make love, not war'" (San Francisco, Norwegen 1997)

• Eskapismus – Sehnsucht nach Ausstieg, Idylle, heiler Welt

"I wish I were an astronaut / Who looks down to the earth from up there / Because all those, who already were up there / They all say it's wonderful / Here you'll feel time slower passing by / Silence and peace that we miss so much" (Keine Grenzen-Zadnych granic, Polen 2003)

"When I look at the hatred there is / The pain, the poverty / I see myself again as a child / Running in the blue / It was wonderful to fly away / to challenge the wind, the imagination" (It was wonderful, Schweiz 2008)

#### Stimulierung von Betroffenheit

Emotionalisierte Darstellung von Unrecht und Leid bzw. des Empfindens des Ich-Erzählers angesichts der Not in der Welt

"I slowly walk into the night around / To see how dreams of people die / ... / And I'm so sorry I'm so helpless in this angry world / If only I could change it for one day" (The war is not over, Lettland 2005)

"And once again boats sail down the Danube / Don't worry about me anymore / I'm like leaves, the wind blows me away / Wolves die alone" (Wolves die alone, Croatien 1995) "My John is eighteen years old / Pray for him, angels' song / Give us peace, give us peace ..." (Don't ever cry, Kroatien 1993)

• Einbezug der Alltagserfahrung der Rezipienten

"A new day, a new scandal / We see violence in the news / Children are thrown in the battle / For ideals, for profits" (Soldiers of love, Belgien 1987)

"You, who buy your indulgence in tickets / Which may win you TVs and cars / What are you afraid of? / Why are you still in doubt? / Where are you when they ask for your support? / You who speak such fair words of peace" (Where are you? Norwegen 1974) "We build houses to withstand hurricanes / And make ships to sail in every storm / ... / It only just won't work to keep peace" (Peace, Niederlande1993)

#### Kuschelfaktor und Tränendrüse – Abwehr bzw. Kanalisierung von Handlungsbereitschaft

- Fördern der Solidarisierungsbereitschaft mit den "Guten" (Ich-Erzählern) und der emotionalen Verwicklung in ihre Botschaft
  - [z. B. Gleichgewicht aus Bedrohung durch Krieg/Not/das Böse + Betonung eigener Stärke durch die richtige Moral, Überzahl, Durchhaltevermögen etc.]
  - "I'm only a little soldier with no weapon but my voice / But I'll cry for you, my dove for you / I'm only a little soldier, now and over there / But a thousand voices will shout it with me" (It takes time (I will fight for that), Frankreich 2002)
  - "Ich weiß, meine Lieder, die ändern nicht viel / Ich bin nur ein Mädchen, das sagt, was es fühlt / allein bin ich hilflos ..." (Ein bißchen Frieden, Deutschland 1982)
  - "Can one voice really make a difference? / Still I know we've got to try" (Why, Finnland 2005)
- Fördern selbstgefälliger Klage und unspezifischer Empörung gegenüber Krieg und Unrecht, die sich in vagem Protest ohne konkrete Adresse erschöpft
  - [z. B. Vage Opfer-/Unrechtsdarstellung + Demonstration eigenen Mitgefühls + Abwesenheit von Verursachern]
  - "Wer gibt den Kindern ihr Lachen zurück / Bevor ihre Hoffnung zerbricht?" "Für die Menschen, die frier'n, und die keiner mehr hört / Dafür singe ich euch dieses Lied" (Für den Frieden in der Welt, Österreich 1994)
  - "I've fought for that, I've clasped in my arms / All those crying women that we can't see" (It takes time (I will fight for that), Frankreich 2002)

#### 5. Soziale Identifikation

#### Identifikation mit einer "Wir-Gruppe

- Anreize zu Identifikation mit der imaginierten Gemeinschaft des Publikums "Everybody, shout: peace will come"
  - "Hi, hi, come on, let's go to Jerusalem all together / We walk hand in hand to a peaceful land (...) Hello, hello, we'll raise out hands with you / Peace to the world from Jerusalem" (Reise nach Jerusalem Kudüs'e seyahat, Deutschland 1999)
- Anreize zur sozialen Identifikation mit dem Ich-/Wir-Erzähler (durch Leid, Sympathie/Attraktivität, die korrekte Botschaft)
  - "We see violence in the news / (...) / We feel guilty / But sip our wine / You're asking yourself constantly / What we are doing" (Soldiers of Love, Belgien 1987)
  - "Wie eine Blume am Winterbeginn / Und so wie ein Feuer im eisigen Wind / Wie eine Puppe, die keiner mehr mag / Fühl ich mich an manchem Tag" (Ein bißchen Frieden, Deutschland 1982)
- Aneignung von Friedens-Ikonen zugunsten des eigenen Anliegens
  - "All over the earth, Lennon's song was heard: / Love, love, love" (San Francisco, Norwegen 1997)
  - "Mr Ghandi has died, will he be dead for a long time?" (It takes time (I will fight for that), Frankreich 2002)
  - "I had a dream where I saw / The whole world with a flag / On which you could read in big: / We want peace and no hatred" (Anspielung auf Martin Luther King's Rede; Soldiers of Love, Belgien 1987)
- Empathie mit "allen" Menschen, unabhängig davon, ob sie Täter oder Opfer sind
  "Kids with guns are always too young to die" (Peace will come, Diana Gurtskaya 2008)
  "We are the children of the whole world / asking for a simple happy life" (Love Card,
  Marokko 1980)

#### Beschränkung von Anreizen zur sozialen Identifikation mit den Opfern/Leidenden

- ... auf bewährte Opfergruppen
  - "Wer gibt den Kindern ihr Lachen zurück, bevor ihre Hoffnung zerbricht?" (Für den Frieden in der Welt, Österreich 1994)
  - "All those crying women that we can't see" (It takes time (I will fight for that), Frankreich 2002)
  - Innocent children die" (Why angels cry, Zypern 2006)
- ... auf vage Leidensbeschreibungen/-klischees
  - "The people lost in fear" (Peace will come, Georgia 2008);
  - "Menschen von Tränen fast blind" "Menschen, die frier'n, und die keiner mehr hört" "Flehende Herzen im Dunkel der Nacht" (Für den Frieden in der Welt, Österreich 1994)

#### Dehumanisierung der Gegner

- Dehumanisierung von Tätern/Kriegführenden, so dass Ursachen verschleiert werden "A love card with warm greetings / Not sent by the traitors of peace" (Love Card, Marokko 1980)
  - "Dark are the misery, the men and the war / Which believe that they hold the time's reins" (The bird and the child, Frankreich 1977)
- · Opfer ohne Täter
  - "We see violence in the news / Children are thrown in the battle / For ideals, for profits" (Soldiers of Love, Belgien 1987)

# 4 Diskussion der Ergebnisse und Ausblick

Angesichts der Gesamtzahl von Liedern, die im ESC seit 1956 jährlich präsentiert wurden, nehmen die Themen Krieg und Frieden in den Texten eine überraschend untergeordnete Stellung ein. Allerdings haben sie seit ca. den 1990er-Jahren einen festen Platz erobert und sind fast in jedem Wettbewerb in mindestens einem Song vertreten. Gerade weil der ESC seinen Teilnehmerländern einen spielerischen, kreativen Rahmen dafür bietet, sich der internationalen (Fernseh-)Öffentlichkeit vorzustellen und sich kulturell in einem erweiterten Europa zu verorten, liegt in einem solchen Format auch die Chance, gemeinschaftliche zentrale Anliegen wie ein friedliches Zusammenleben auf die Agenda zu setzen. Wir gingen der Frage nach, wie sich über Frieden und Krieg singen lässt, wenn deren politische Dimension durch regulierende Vorschriften ausdrücklich ausgeklammert wird, und fanden in den Songtexten einen überaus starken Gemeinschaftsbezug. Dieser erschöpft sich größtenteils im selbstreferenziellen Feiern der eigenen Friedfertigkeit, im kollektiven Beklagen einer unvollkommenen und im gemeinsamen Erträumen einer besseren Welt. Wir waren neugierig auf die Umsetzung von Merkmalen einer Deeskalationsorientierung und der Zurückweisung von Eskalationsorientierung. Wir fanden etwas gänzlich anderes, nämlich Kitsch.

Unser anfänglicher Versuch, das kempfsche Modell ohne Modifikationen auf Schlagertexte anzuwenden, könnte als naiv bezeichnet werden. Dafür sprach jedoch das Prinzip der Einfachheit, welches sinngemäß besagt, dass bei konkurrierenden Theorien diejenige vorzuziehen ist, die mit weniger Annahmen auskommt. Anders gesagt: Wir wollten das kempfsche Modell an dieser anderen Textsorte einfach einmal ausprobieren, auch im Sinne von auf die Probe stellen oder testen. Nach den ersten Kodierversuchen wurde klar, dass die Analyse von Schlagertexten eine Anpassung des Analyseinstruments erfordert, da diese Texte häufig auf eine Weise inhaltlich und stilistisch Emotionen ansprachen, wie man es am ehesten im Kitsch findet.

Im Unterschied zum Modell von Kempf fehlt in unserer Arbeit der Gegenpol – Kriegskitsch. Dieser kann gewiss aufgrund theoretischer Überlegungen, insbesondere Analogiebildung konstruiert werden. Für eine empirische Ableitung des Pols boten die ESC-Songs (mit Ausnahme von "Dschingis Khan") jedoch so gut wie kein Material. Dies führen wir in erster Linie auf das Format eines internationalen Schlagerfestivals zurück. Wir sind überzeugt davon, dass es möglich ist, empirische Beispiele für Kriegskitsch in Songtexten anderer Musikstile wie etwa Punk, Heavy-Death- und Mittelalter-Metal oder Gothic, möglicherweise sogar auch im Volkslied, zu finden und das Instrument mit dem Fokus auf eine Romantisierung von Krieg und Gewalt, auf die Verklärung von Kriegshelden etc. zu ergänzen.

Als Ergebnis können wir schließlich "nur" ein in den Daten gründendes, vorläufiges Instrument anbieten, die *Checkliste Friedenskitsch*, aber noch keine statistisch abgesicherten Ergebnisse der Anwendung dieses Instruments. Wir betrachten unsere Checkliste als stellenweise noch skizzenhaften Entwurf, den empirisch anzuwenden und weiter auszuarbeiten sich Johnen würde. Insbesondere die

Erweiterung durch Aspekte der Motivationslogik, der Harmonisierung der Bezugsebenen und zweiseitige Botschaften steht noch aus. Gerade hinsichtlich mythischer Bezüge bieten die von uns gefundenen "Friedens"-Texte eine unterhaltsame Fundgrube. Unseren Entwurf möchten wir nun interessierten Kolleg(inn)en oder Studierenden zur weiteren Verwendung und Überarbeitung oder bereits zur Analyse überlassen.

Interessant wäre mit Sicherheit auch die Untersuchung von Friedenskitsch nicht nur in Songtexten, sondern in der Kombination aus Text, Musik und der visuellen und auditiven Inszenierung der Stücke. Dies war von uns ursprünglich ebenfalls angedacht, hätte aber den Rahmen unseres Aufsatzes (und unsere eigenen Kapazitäten) gesprengt. Einer Untersuchung wert wäre außerdem die Anwendung des Instrumentes auf Songtexte aus anderen Kontexten, außerhalb des ESC, die keinerlei Einschränkungen durch eine "apolitische" Klausel unterliegen. Blüht die Blume des Kitsches nur unter den Bedingungen eines Tabus expliziter Aussagen zu Krieg oder Frieden? Wie wird außerhalb des ESC mit diesen Themen umgegangen?

Auch die grundsätzliche Frage nach Funktion, Produktionsweisen und Bedingtheiten von "Friedenstexten" in einem Unterhaltungsformat wie dem ESC und in einem Unterhaltungsmedium wie dem öffentlich-rechtlichen Fernsehen haben wir ausgeblendet. Möglicherweise ist bereits das Medium die ganze Botschaft (McLuhan & Fiore 1984). Frieden oder Krieg wären dann ebenso wie Liebe, Glück, Einsamkeit oder Schmerz willkürlich gewählte, größte gemeinsame Nenner aus dem Bauchladen menschlicher Grunderfahrungen, mit denen sich ein vorgestelltes internationales Massenpublikum in irgendeiner Weise identifizieren können soll. Welche Produktions- und Selektionsprozesse im Vorfeld des ESC auf welchen Ebenen stattgefunden haben und unter welchen Bedingungen ein so zentrales Thema wie Frieden hierfür funktionalisiert wird, wäre Gegenstand einer ganz anderen Untersuchung. Hier haben wir uns hingegen bewusst auf die Textanalyse bzw. die Konstruktion eines Analyseinstrumentes beschränkt, für das die ESC-Songtexte das Rohmaterial boten.

Dass eine Checkliste Friedenskitsch auch für die Untersuchung der Presseberichterstattung sinnvoll und nützlich sein könnte, daran besteht für uns kein Zweifel. Bei der qualitativen Untersuchung der deutschen Nachkriegsberichterstattung über Frankreich (Jaeger 2009) hätte sie bereits wertvolle Dienste leisten können, zumal ausgerechnet einige der besonders "deeskalationsorientierten" Texte auch besonders idealisierend und rührselig eine deutsch-französische Romanze heraufbeschworen. Mehr als einmal kam der Verdacht auf, es gehe vor allem darum, der deutschen Leserschaft ein wohliges Gefühl der Ergriffenheit angesichts ihrer eigenen, endlich errungenen Läuterung und Friedensliebe zu verschaffen.

<sup>7</sup> Z. B. steckt in "Ein bißchen Frieden" der Mythos des unschuldigen Kindes, das durch sein reines gutes Herz und seinen Gesang die Welt erlösen kann; in "Humanahum" gibt es einen "alten Weisen", der Kindern vom Verlust des Paradieses erzählt; in "It takes time (I will fight for that)" nimmt es David ("I'm just a little soldier with no weapon but my voice") mit einem übermächtigen Goliath (unbestimmt durch Krieg, Gewalt und Unrecht) auf.

Friedenskitsch, wie wir ihn in seinen verschiedenen Ebenen und Ausprägungen bzw. Variablen und Beispielen definieren, ist letztlich eine Abwehrfigur, die eine angemessene Reaktion auf und die ernsthafte Beschäftigung mit Krieg und Frieden sowie den eigenen diesbezüglichen Handlungsmöglichkeiten verhindert. Die Thematisierung von Krieg und Frieden in der Unterhaltungsindustrie weist daher strukturell große Ähnlichkeit mit der im Eingangszitat beschriebenen Doppelbindungssituation im Golfkrieg auf, das wir wie folgt modifizieren können:

Der Widerspruch zwischen maßloser Medienpräsenz und gleichzeitig lückenlosem Kitsch verwickelt das Publikum in eine Situation, die alle Definitionsmerkmale der Doppelbindung aufweist: infolge ihres Unterhaltungsmonopols besteht eine intensive Abhängigkeit von den Medien; mit dem Anspruch, durch die Thematisierung von 'Krieg und Frieden' auch realitätshaltige Unterhaltung zu liefern, bei gleichzeitig kitschiger Form aller Inhalte, übermitteln die Medien jedoch zwei widersprüchliche Botschaften, zu denen man wegen des Mangels an unabhängigen Informationsquellen nicht Stellung beziehen, sich aber infolge der Allgegenwart der Unterhaltungsindustrie auch nicht zurückziehen kann.

# Mediendiskurs, Gewalt und gewaltfreie Konfliktbearbeitung. Das Journalism-in-the-New-World-Order-Projekt oder: von der Faszination einer Methode

Stig Arne Nohrstedt & Rune Ottosen

# 1 Die Vorgeschichte

Die Idee, ein internationales Forschungsprojekt über Journalism in the New World Order zu initiieren, kam im Sommer 1991 auf, als wir an einer IAMCR-Tagung über den Golfkrieg (1990-91) in Istanbul teilnahmen. Damals waren die theoretischen und politischen Diskurse über die Globalisierung im Allgemeinen wie auch in den Medien im Besonderen recht abgehoben und flogen hoch in dünner Luft auf abstrakten und euphorischen Flügeln. Wir bedauerten das Fehlen solider empirischer Studien und kamen zu dem Schluss, dass ein globales Netzwerk von Forschern die conditio sine qua non sei, um die Folgen der Globalisierung studieren zu können. Wir versuchten Geld aufzutreiben und begannen gleichzeitig, nach Kollegen zu suchen, die sowohl an diesem Projekt interessiert waren als auch ergänzende methodische Kompetenzen einbringen konnten.

Wir hatten das Glück, Wilhelm Kempf kennenzulernen, einen deutschen Wissenschaftler und Latent-Class-Analyse-Spezialisten, der aus der Psychologie kommt und der damals der Vorsitzende des Forums Friedenspsychologie war – er bündelte mehr Kompetenzen für das künftige Projekt, als wir je zu träumen gewagt hätten. Damit begann nicht nur ein sehr erfolgreiches internationales Forschungsprojekt, sondern auch eine langjährige kollegiale Beziehung und Freundschaft.

Wir haben keinen Zweifel daran, dass das internationale Medienforschungsprojekt Journalism in the New World Order (NWO), das wir gemeinsam aus der Taufe hoben, für die Kriegsberichterstattungsforschung sehr bedeutsam war. Auch uns haben die Fragen, die sich uns damals stellten, bis heute, zwanzig Jahre danach, immer weiter beschäftigt. Wie Wilhelm Kempf haben auch wir uns nicht nur mit fehlerbehafteter und manipulierter Berichterstattung auseinandergesetzt, sondern ebenso nach Alternativen dazu gesucht. Eine dieser Alternativen ist die Richtung, die Johan Galtung (2002) Friedensjournalismus genannt hat. 2011 sehen wir für uns selbst wie auch für alle Kollegen eine dringende Aufgabe darin, sich damit zu beschäftigen, wie sich die Kriegsberichterstattung nach dem Ende des Kalten Krieges verändert hat. Nach dem Fall der Berliner Mauer 1989 und der Auflösung der

Sowjetunion 1991 sollte eine neue Weltordnung errichtet werden, in der internationale Gerechtigkeit und Frieden herrschen würden. Jedoch die Kriege auf dem Balkan in den 1990er-Jahren ließen ein Konzept der Neuen Kriege entstehen, das nach den Terroranschlägen des 11.09.2001 eine Bildsprache erlangt hat, die der von Spielfilmen den Rang abläuft. Der Golfkrieg 1990/91 brachte den kommerziellen Durchbruch für den rund um die Uhr berichtenden Nachrichtensender CNN und der Krieg in Afghanistan 2001 für seinen Rivalen al-Jazeera. Im Irakkrieg 2003 schlug die große Stunde des Internet. Salam Pax ist das Pseudonym eines Bloggers aus Bagdad, der zur ersten Generation der Kriegsblogger gehört. Im Juli und Dezember 2010 veröffentlichte WikiLeaks mehr als 300,000 Geheimdokumente über den Afghanistan-Krieg und diplomatische Berichte, die eine Fülle von Berichten über den Tod von Zivilisten und Enthüllungsgeschichten illegaler Operationen auch in den Mainstream-Medien nach sich zogen. Eine neue Weltordnung, neue Kriege, neue Medien – welchen Einfluss hat all das auf die Kriegsberichterstattung? Können wir Anzeichen einer neuen Kriegsberichterstattung ausmachen, vielleicht sogar die Entwicklung des Friedensjournalismus, den Wilhelm Kempf später zu einem friedenswissenschaftlichen Programm (vgl. Kempf 2007) weiterentwickelt hat?

Auch wenn es einige aktuelle Forschungsschwerpunkte gibt, die unserer Aufmerksamkeit wert wären, wollen wir im Folgenden die Perspektive verengen und das Design, die Methodologie und die Ergebnisse des NWO-Projekts vorstellen.

# 2 Zur Entstehung des NWO-Forschungsprojekts

# 2.1 Politisches Klima und Hintergründe

Nach dem Fall der Berliner Mauer 1989 gab es große Hoffnungen auf eine Welt mit friedlichen, engeren Beziehungen zwischen den Ländern, den Kontinenten und den Kulturen. Die Menschheit freute sich auf eine bessere und glücklichere Zukunft, auf eine neue internationale Ordnung, in der die Kommunikation zwischen verschiedenen, z. T. weit voneinander entfernten Teilen der Welt sich rapide vermehren würde, was zum gegenseitigen Verständnis beitragen würde. Globalisierung wurde das Schlagwort der 90er-Jahre des 20. Jahrhunderts, die magische Formel für Versprechungen einer sich verkleinernden Welt, in der sich Raum und Zeit verdichten und in der sich auch das vorhergesagte global village tatsächlich entwickeln kann (vgl. Giddens 1990; McLuhan 1964).

Die Welt war dabei, "a single place" (Robertson 1992) zu werden. Eine Folge dessen, dass dort dann alle Menschen das gleiche Schicksal teilen würden, sah man darin, dass sie ihre Auseinandersetzungen notwendigerweise mit friedlichen Mitteln austragen müssten. Es hatte den Anschein, als würde sich die Utopie, die in der UNESCO-Debatte der 1970er-Jahre um eine Neue Weltinformationsordnung (NWICO) zum Ausdruck gekommen war, mit all ihren hochfliegenden Hoffnungen auf weltweite Kommunikation ohne die Dominanz von Gewalt und Ungleichheit letztlich doch noch erfüllen (Nohrstedt & Ottosen 2010).

Bald jedoch wurde die Euphorie nach dem Fall der Berliner Mauer von der Realität eingeholt, und zwar in Gestalt einer Kriegspropaganda, die alle utopischen Erwartungen zunichte machte. Als US-Präsident George H. Bush 1990 im In- und Ausland begann, Stimmung zugunsten einer militärischen Intervention im Persischen Golf zu machen, um die irakische Besetzung Kuwaits zu beenden, so geschah dies unter dem Versprechen einer Neuen Weltordnung. Internationales Recht und Frieden würden sich letztlich durchsetzen und die Zukunft der Weltgemeinschaft bestimmen. Nun sei die Zeit gekommen, in der Diktatoren nicht länger Erfolg damit haben würden, ihre Nachbarländer zu bedrohen und zu besetzen, wie es Saddam Hussein im August 1990 getan hatte. Solche Hoffnungen, hier übersetzt in Kriegspropaganda, widerspiegelte auch Francis Fukuvamas These vom Ende der Geschichte, derzufolge die Zeit nach dem Kalten Krieg gekennzeichnet sein sollte von pragmatischer, friedlicher Politik und nicht von fundamentalen Konflikten um die Weltherrschaft zwischen konkurrierenden politisch-ideologischen Systemen (Fukuyama 1989). Aber es zeigte sich, dass der Golfkrieg – ein Krieg, der Baudrillard (1991) zufolge "niemals stattfand" – alles andere war als die saubere chirurgische Operation, wie dies von den Alliierten behauptet worden war. Und auch wenn es für einen kurzen historischen Moment so ausgesehen hatte, als ob die internationale Gemeinschaft wie ein Mann gegen den Aggressor aufstehen würde, so dauerte es doch gar nicht lange, bis die These vom Ende der Geschichte durch die neue vom Kampf der Kulturen (Huntington 1996) ersetzt wurde.

## 2.2 Der Golfkrieg und seine Folgen

Die Ergebnisse unserer Irak-Studie deuten darauf hin, dass die Bedeutung des Golfkriegs als globales Medienereignis, das die Entstehung zukünftiger globaler Konflikte beeinflussen wird, leider – und in weitaus höherem Maße als wir denken – unterschätzt wurde; dass uns der Golfkrieg noch Jahrzehnte, wenn nicht sogar Jahrhunderte lang verfolgen wird (Nohrstedt & Ottosen 2001, Kempf & Luostarinen 2002). Das ist der Tatsache geschuldet, dass in diesem Krieg eine Auseinandersetzung zwischen der muslimischen und der westlichen Welt stattfand. Viele Muslime betrachteten den Krieg als Angriff auf die muslimische Nation (al-Umah). Es war unter ihnen umstritten, dass arabische Staaten wie Saudi-Arabien der von der UNO unterstützten Militärallianz beitraten, um die irakischen Streitkräfte aus Kuwait zu jagen. Deshalb konnte Saudi-Arabien danach zur Brutstätte für radikalislamistische Gruppierungen und Terrororganisationen wie al-Qaida werden. Gleichzeitig gelang im Golfkrieg modernen Propagandatechniken der globale Durchbruch, z. B. vermittelt durch Direktübertragungen von CNN.

In unserer vergleichenden Studie konnten wir herausarbeiten, dass die freiwillige Unterwerfung der Medien unter die amerikanische Lesart in ihrer Berichterstattung eine große Rolle beim Framing der Golfkriegsgeschichte spielte (Nohrstedt & Ottosen 2001). Wenn wir dies mit der weltweiten Berichterstattung über den 11. September vergleichen, stellen wir in vielen Ländern (auch in Norwegen) einen nationalen Fokus auf den lokalen Bedrohungsaspekt als gemeinsames Merkmal fest (Ottosen & Figenschou 2007). Die Medialisierung der Bedrohungspolitik wie auch

der Politik überhaupt impliziert, dass sowohl die Medien als auch die Journalisten immer mehr der Kultur der Angst unterliegen und diese auch zunehmend weitertragen (Nohrstedt 2011). Sowohl Terroristen als auch ihre Gegner wenden im weltweiten Krieg gegen den Terror bestimmte Medienstrategien an, um die Herzen und Köpfe der Öffentlichkeit für sich zu gewinnen. Jedoch mangelt es den Journalisten und den Medienunternehmen grundsätzlich an effektiven Gegenstrategien gegen Versuche, die Sensibilität der Nachrichtenmedien für spektakuläre, dramatische und bedrohliche Nachrichten und entsprechende Visualisierungen auszunutzen.

## 2.3 Die Vernetzung der Forscher

Um medialisierte Globalisierungsprozesse erfassen zu können, war es notwendig, nicht nur die Medien verschiedener Länder, Kontinente und Kulturen zu untersuchen, sondern auch Wissenschaftler unterschiedlicher Forschungstraditionen und Schulen in das Projekt einzubeziehen. Deshalb schätzten wir uns glücklich, nicht nur Wilhelm Kempf und die Studenten seiner Forschungsgruppe für das Projekt interessieren zu können, sondern auch Heikki Luostarinen, unseren finnischen Kooperationspartner an der Universität Tampere, Hamid Mowlana von American University in Washington D.C., Oddgeir Tveiten, (damals) an der Universität Bergen; Hanne Mathiesen, eine Forschungsassistentin am Internationalen Friedensforschungsinstitut Oslo und Admassu Tassew, einen Ph.D.-Studenten aus Äthiopien an der Universität Göteborg in Schweden. Auch beteiligten sich andere Kollegen und Studenten unserer Heimatinstitutionen und investierten viele Arbeitsstunden, unterstützten uns mit wertvollen Kommentaren und Kritik. Sicherlich führte selbst dieses Netzwerk von Forschern nicht zu einer wirklich globalen Analyseperspektive, doch es war mindestens ein erster Schritt in eine Richtung, die später zu einem weltweiten Projekt ausgebaut wurde (vgl. z. B. Nohrstedt & Ottosen 2004, 2005). Wir können uns nur wünschen, dass künftige Projekte weitere Schritte unternehmen werden, um den Pfad der viel zu verbreiteten verwestlichten Perspektiven auf den Globalisierungsprozess zu verlassen.

# 3 Projektdesign: Theoretische Vorannahmen – methodisches Vorgehen – Forschungsergebnisse

Unser Projekt war sowohl multidisziplinär als auch multi-methodologisch angelegt. Seine grundlegende empirische Basis bildet die quantitative Inhaltsanalyse einer Stichprobe von ungefähr 4.000 Pressemeldungen in Qualitätszeitungen und im Fernsehen aus sieben Ländern: aus Äthiopien, Finnland, Deutschland, dem Irak, Norwegen, Schweden und den USA. Aus praktischen Gründen waren die Stichproben aus Äthiopien und dem Irak relativ begrenzt. Die anderen aber reichten aus, um eine vergleichende Analyse der nationalen Framings in der Berichterstattung über den Golfkrieg durchführen zu können. Diese größeren Stichproben aus Europa und den USA wurden außerdem Latent-Class-Analysen unterzogen, was das Projekt einzigartig in dem Sinne machte, dass in länderübergreifenden verglei-

chenden Medienstudien komplementäre Methoden angewandt wurden. Für die kleineren Samples wurde auch eine Variante der von den Forschungen Norman Faircloughs angeregten kritischen Diskursanalyse durchgeführt (Nohrstedt & Tveiten 2002).

# 3.1 Die Latent-Class-Analyse

Der wichtigste Beitrag, den Wilhelm Kempf zu dem Projekt geleistet hat, besteht darin, dass er die Latent-Class-Analyse (LCA) als Instrument auf die Analyse von Medieninhalten übertragen hat. Diese Methode kombiniert quantitatives und qualitatives Herangehen und ermöglicht es, latente Klassen typischer Berichterstatungsmuster, welche Kempf Stile nennt, zu identifizieren. Die Stile sind durch spezifische Wahrscheinlichkeiten charakterisiert, mit denen (auf einen Konflikt bezogene) Variablen in ihnen vorkommen, und sie repräsentieren bestimmte Standpunkte gegenüber dem Konflikt. Die LCA zeigt, wie häufig die verschiedenen Stile in einem bestimmten Medientyp oder auch in der Medienlandschaft eines bestimmten Landes vorkommen und wie sie miteinander kombiniert sind.

Wir konzentrieren uns im Folgenden auf die kurze Zusammenfassung einiger Ergebnisse der LCA (vgl. Nohrstedt & Ottosen 2001). Kempf et al. thematisieren in ihrem Beitrag *New World Order Rhetoric in US and European Media* die zentrale Frage für das Gesamtprojekt, und zwar die Frage danach, welche Rolle die Vision einer neuen Weltordnung in der Propaganda wie auch in den Medienberichten spielte. Sie untersuchten die Berichterstattungsstile, die in den Nachrichten und Editorials identifiziert werden konnten, daraufhin, ob Schlüsselelemente der US-Propaganda hinsichtlich der Intervention im Irak darin vorkommen. Folgende rhetorische Muster bzw. Stile konnten identifiziert werden:

- Lehren aus der Geschichte Lehren aus der Vergangenheit: Diktatoren dürfen nicht ermutigt werden durch Anzeichen von Schwäche; und/oder ein Bezug zur Geschichte wird hergestellt (Beschwichtigungspolitik/Hitler), um diese Schlussfolgerung nahezulegen.
- Der rechte Moment Wo stehen wir jetzt. Der Augenblick für einen Neuanfang darf nicht verpasst werden; und/oder ein Bezug zur gegenwärtigen Situation wird hergestellt (Ende des Kommunismus/Sieg der Demokratie/Einsatz der UNO), um diese Schlussfolgerung nahezulegen.
- 3. Fair play Wovon werden wir momentan geleitet: Das Ziel der neuen Politik besteht darin, Gerechtigkeit und ethische Prinzipien zu Leitlinien der internationalen Beziehungen zu machen; und/oder es wird auf die Rechte kleiner Nationen verwiesen, um darauf hinzuweisen, dass die Früchte der Gewalt nicht hingenommen werden dürfen usw. (Kempf et al. 2001, S. 129f.).

Diese Stilmittel einer Rhetorik der Neuen Weltordnung wurden in den über 4.000 Nachrichten nur 136 Mal identifiziert, in den Leitartikeln bzw. Kommentaren aber viel häufiger (187 Mal). Eine anschließende gemeinsame Analyse von Leitartikeln und Nachrichten ergab jedoch ein konsistentes Bild. Wenn die Rhetorik der Neuen Weltordnung erwähnt wurde, fand sie in 90% der Fälle Zustimmung. Die Stile 1

und 3 wurden am häufigsten verwendet. Lehren aus den Erfahrungen des Zweiten Weltkriegs zu ziehen ist von fast ebenso überragender Bedeutung wie die Hoffnung, dass die Neue Weltordnung eine Zukunft ermöglichen wird, in der Gerechtigkeit, internationales Recht und Menschenrechte herrschen werden. Jetzt sei auch der richtige Zeitpunkt dafür, dies auf den Weg zu bringen. Im Unterschied zur Verwendung anderer Methoden, insbesondere reiner Häufigkeitsauszählungen, ermöglichte es die LCA, bei der Verwendung des Stils Lehren aus der Vergangenheit eine Rangordnung der Medien der verschiedenen Länder aufzustellen: Dieser Stil kommt am häufigsten in den norwegischen Medien vor, gefolgt von den USA, danach Deutschland, Finnland und Schweden.

Im seinem Beitrag *News Media and Conflict Escalation* fokussiert Kempf, wie die Medien Konflikte und deren Hintergründe üblicherweise konstruieren und zeigt Wege auf, zu anderen Darstellungen zu gelangen. Damit stellt der Beitrag nicht nur wertvolle empirische Ergebnisse zur Verfügung, sondern ist auch selbst ein wichtiges Stück Friedensjournalismus-Literatur. Einleitend erläutert Kempf et al. (ebd., S. 149):

"The current discussion among journalists and scholars in media research and peace studies focuses on how the media can become vehicles for conflict prevention and for constructive, nonviolent, conflict transformation, including the problem of reporting on how to repair wartorn societies – in material, social and human terms. Proposing a critical peace journalism as an alternative to traditional war reporting does not mean suggesting that violence should not be reported, or that what are reported as facts should not be empirically correct. What it does stress, however, is that the *form* in which they are reported may contribute either to the escalation or to the de-escalation of the conflict."

Folgende Berichterstattungsstile konnten zur Beschreibung des Konflikthintergrunds identifiziert werden:

- 1. das Interesse des Irak, die Ölvorkommen zu kontrollieren und/oder seine Macht und Vorherrschaft zu stärken.
- das Interesse der Alliierten (oder einiger von ihnen), die Ölvorkommen zu kontrollieren und/oder ihre Macht und Vorherrschaft zu stärken,
- 3. der Anspruch des Irak auf historische Rechte an kuwaitischem Territorium,
- Kuwaits Weigerung, mit dem Irak über die Grenzstreitigkeiten und/oder über die Höhe der Ölexporte zu verhandeln und
- 5. die irreführende Diplomatie der USA (oder anderer Alliierter).

# 3.2 Ergebnisse

Die Ergebnisse der LCA der Medienberichterstattung in den nordischen Ländern waren für unsere eigenen Forschungen und Veröffentlichungen außerordentlich fruchtbar (Nohrstedt & Ottosen 2001, 2004, 2005, 2010). Wie Kempf (2002) herausarbeitet, unternahmen die finnischen Medien die geringsten Anstrengungen, den Konflikt zu erklären. Die norwegischen Medien folgten dem allgemeinen Muster am ehesten, indem sie konstatierten, der Irak sei schuld. Dazu befanden sich die schwedischen Medien in relativer Distanz. Dies ist wesentlich für die Ergebnisse

des NWO-Projekts insgesamt. Von herausragender Bedeutung ist dabei die Tatsache, dass die Konfliktberichterstattung der Medien an die Sicherheitspolitik des jeweiligen Landes angepasst ist. Das NATO-Mitglied Norwegen folgt dem Stil der US-Medien deutlicher als das formal blockfreie Schweden. Außerdem konnte Kempf nachweisen, dass die schwedischen Medien dazu neigten, die historischen Ansprüche des Irak, die Abhängigkeit der Alliierten von Öllieferungen und deren Hegemonialinteressen häufiger zu erwähnen als die Medien fast aller anderen Länder (ebd., S. 170f.).

Über die Motive der Alliierten schwiegen sich die skandinavischen Medien recht deutlich aus. Dies implizierte auch die Befreiung Kuwaits, das Hauptpropagandathema der USA. Die schwedischen Medien, die das größte Stillschweigen wahrten, bemühten sich, den Eindruck zu vermeiden, dass das Interesse der Alliierten an diesem Krieg in ihren Ölinteressen liegen könnte (ebd., S. 153). Die Unterschiede zwischen den schwedischen und den norwegischen Medien schlugen sich auch deutlich darin nieder, wie sie die Rolle der UNO bestimmten:

"The UN role was also clearly important in the Swedish media, where the UN was idealized (more often than in all other countries) as an instrument for peace (style 3) and (nearly as often as in Finland) as an instrument for the enforcement of international law (style 4). Though the UN role plays a relatively smaller role in the Finnish media, there is a tendency similar to that of the Swedish media: the UN was idealized (nearly as often as in Sweden) as an instrument for peace (style 3) and (more often than in any other country) as an instrument for the reinforcement of international law (style 4)." (ebd., S. 159)

Die norwegischen Medien schenkten der Rolle der UNO fast so wenig Aufmerksamkeit wie die der USA, wohingegen die finnischen Medien der UNO und den Initiativen Dritter Parteien beträchtliche Aufmerksamkeit widmeten. Eine vergleichende Diskussion friedlicher Alternativen wurde von ihnen viel öfter als in allen anderen westlichen Ländern angeboten (mit Ausnahme Deutschlands). Auch die schwedischen Medien beschäftigten sich mit diesem Thema, und zwar meist im Zusammenhang mit Initiativen Dritter Parteien und der UNO. Darüber hinaus fokussierten sie die ökonomischen Sanktionen stärker als alle anderen. Die norwegischen Medien waren am wenigsten bereit, über Möglichkeiten einer friedlichen Konfliktbeilegung zu diskutieren. Initiativen dritter Parteien und alternative Lösungen wurden – ebenso wie von den US-amerikanischen Medien – eher selten erwähnt (ebd., S. 163).

Die norwegischen Medien berichteten sehr einseitig über Verhandlungsverweigerungen und wiesen die Schuld daran ausschließlich dem Irak zu. Ebenso wie die anderen skandinavischen Medien machten sie eher keine Aussagen sowohl über die eigentlichen Interessen der Alliierten als auch über das Ziel, Kuwait zu befreien. Besonders selten wurde darüber berichtet, wie die USA und ihre Verbündeten im Namen der UNO agieren bzw. dass die UNO von den USA und/oder Großbritannien dominiert sei. Dies könnte darauf hindeuten, dass die norwegischen Medien versuchten, potenziell USA-kritische Argumente zu vermeiden.

Andererseits sind die norwegischen Medien von allen nordischen Medien diejenigen, die die militärische Logik am seltensten unterstützten. Ebenso wie die USamerikanischen berichteten sie kaum über abgelehnte Verhandlungen, Waffenstillstand und Friedensverhandlungen. Und wenn sie dies taten, dann ging es einseitig um ausschließlich vom Irak verschuldete Verhandlungsverweigerungen (ebd., S. 165f.).

Das schwedische Selbstbild, ein neutrales und friedliebendes Land zu sein, hat eine lange Tradition. Als kleiner blockfreier Staat verfolgt Schweden auch eine klar UNO-freundliche Politik. Wie bereits erwähnt, gingen die schwedischen Medien auf relative Distanz zur üblichen Art der Konfliktdarstellung, nach der allein der Irak die Schuld trägt. Öfter als die Medien der meisten anderen Länder wiesen sie auf die Gebietsforderungen des Irak und auf die Abhängigkeit der Alliierten von Öllieferungen hin. Von allen skandinavischen Medien schwiegen sie sich am deutlichsten über das Interesse der Alliierten am Krieg aus. Ganz besonders schien ihnen daran zu liegen, den Eindruck zu vermeiden, dass die Ursache der Intervention in den Ölinteressen der Alliierten liegen könnte.

Auf der anderen Seite wird deutlich, dass die Rolle der UNO sehr wichtig für die schwedischen Medien war. Häufiger als in allen anderen Ländern wurde die UNO idealisiert sowohl als Friedensbringerin als auch (fast ebenso oft wie in Finnland) als Mittel zur Durchsetzung internationalen Rechts. Das Thema friedliche Konfliktbeilegung wurde oft im Zusammenhang mit Initiativen dritter Parteien und der UNO behandelt. Darüber hinaus thematisierten die schwedischen Medien ökonomische Sanktionen am häufigsten. Und wie in den anderen skandinavischen Medien auch fand die militärische Logik kaum Erwähnung (ebd., S. 159).

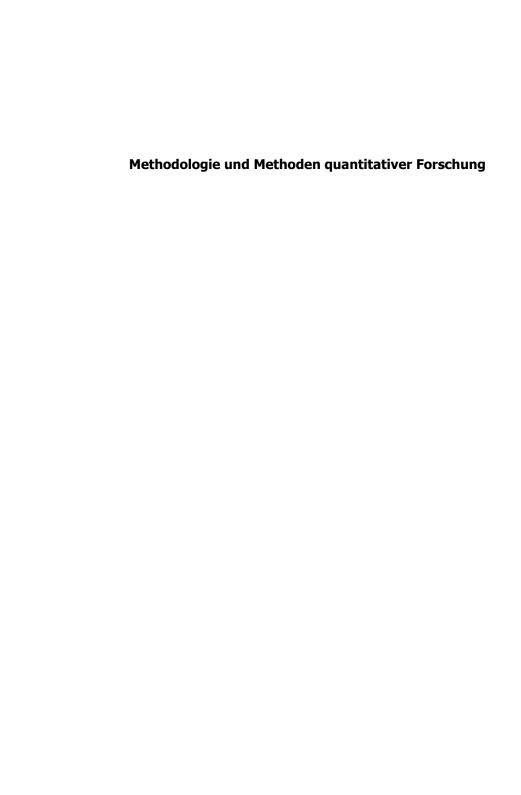
# 3.3 Alternative Möglichkeiten, den Konflikt beizulegen

Die LCA ermöglicht es, den empirischen Nachweis zu erbringen, dass und wie Propaganda die Medienberichterstattung beeinflusst. Im Unterschied zu den deutschen Medien, die ein wenig zum Friedensjournalismus neigten, folgte die Berichterstattung der US-amerikanischen Medien einer strategischen Militärpropaganda. Über die Weigerung der Alliierten zu verhandeln wurde in normalem Ausmaß berichtet, über jegliche (hauptsächlich irakische) Verhandlungsbereitschaft jedoch überhaupt nicht. Noch einseitiger in ihrer Berichterstattung über Verhandlungsinitiativen waren die norwegischen Medien: Selten wurde die (irakische) Verhandlungsbereitschaft wie auch die Verweigerung von Verhandlungen durch die Alliierten thematisiert. Da es für die US-amerikanische Öffentlichkeit offenbar selbstverständlich war, dass man mit dem irakischen Despoten nicht verhandeln kann, war folglich auch Bagdads Verhandlungsbereitschaft nicht von Interesse. Für die norwegische Öffentlichkeit jedoch wäre Norwegens Beteiligung am Krieg nicht akzeptabel gewesen, wenn die irakische Verhandlungsbereitschaft wie auch die Weigerung der Alliierten, nach einer friedlichen Lösung zu suchen, bekannt geworden wären. Der deutliche Bias der norwegischen Berichterstattung zur US-amerikanischen Perspektive wird durch die Ergebnisse der LCA ganz eindeutig belegt (ebd., S. 173f.).

#### 4 Resümee

Das multi-methodologische Design des NWO-Projekts stellte ein großes Potenzial zur Verfügung, das Triangulation und Mehrfachvergleiche ermöglichte. Durch Latent-Class-Analysen gelang es, sogar "versteckte" Berichterstattungsstile deduktiv zu identifizieren und somit deutlich über das hinauszugehen, was mit traditionellen inhaltsanalytischen Methoden möglich gewesen wäre. Wenn überhaupt Kritik an der LCA zu üben wäre, wie sie im Projekt eingesetzt wurde, dann bestünde diese darin, dass es sehr schwierig ist, feine Unterschiede und Veränderungen im Kriegsdiskurs zu entdecken, was ihrer Grundannahme propagandabezogener Stile geschuldet ist. Diese Stile oder Strukturen treten nicht immer sehr häufig in den Nachrichten auf, weshalb das Material für diachrone Analysen nicht ausreicht. Aber dies wird durch die mögliche hohe Präzision und die differenzierte wie auch anspruchsvolle Rekonstruktion von Konfliktdiskursen bei weitem aufgewogen.

Wir haben uns hier auf einige der wichtigen Beiträge konzentriert, die Wilhelm Kempf zusammen mit anderen Projektmitarbeitern und Studenten zum NWO-Projekt geleistet hat. Die Anwendung der LCA in der von Kempf für die Medienforschung weiterentwickelten Form eröffnet völlig neue Perspektiven für Medienstudien und verdient weit mehr Anerkennung als bisher. In unserem Gemeinschaftsprojekt warf die LCA neues Licht darauf, wie Kriegspropaganda die Medienberichterstattung thematisch und rhetorisch beeinflusst. Die gemeinsame Betrachtung der Ergebnisse der LCA und der traditionellen Inhaltsanalyse erlaubt die Identifizierung manifester und systematischer Muster, die bestätigen, dass die Medienberichterstattung über den Golfkrieg 1991 von der Außen- und Sicherheitspolitik der USA beeinflusst war.



# Nichtparametrische spezifisch objektive Modelle für d-dimensionale Testdaten

Hartmann Scheiblechner

# 1 Ein spezifisch objektives eindimensionales Modell für ordinale Daten (ISOP)

Wir nehmen in der klassischen Physik (centimetre, gram, second; CGS-System) als selbstverständlich an, dass – hinreichende Sorgfalt vorausgesetzt – die Messinstrumente beliebig oft reproduzierbare Messergebnisse liefern und die Messergebnisse einer physikalischen Größe (Strecke, Masse, Zeit) nicht vom spezifischen verwendeten Messinstrument abhängig sind. Diese Invarianz hat Rasch spezifische Objektivität genannt. Um wie viel komplexer die klassische Physik wäre, wenn die Messungen nicht spezifisch objektiv wären, kann man an der Relativitätstheorie ermessen.

Die spezifische Objektivität der Messungen ist ein hohes Ziel im Sinne möglichst einfacher empirischer Theorien. In der Physik wird sie durch möglichst exakte Begriffe und Definitionen und exaktes Experimentieren erreicht, in den probabilistischen Modellen von Rasch durch erschöpfende Statistiken und die Technik der bedingten Maximum Likelihood (conditional maximum likelihood, CML-Methode). Diese Methode führt allerdings mit entsprechenden Kontinuitäts- und Differenzierbarkeitsvoraussetzungen dazu, dass nur exponentiale Verteilungsfamilien zugelassen sind. Durch messtheoretische Formulierungen kann man diese unnötig engen Einschränkungen vermeiden und nichtparametrische spezifisch objektive probabilistische Modelle erreichen, die zusätzlich besser und direkt empirisch überprüfbar sind.

Eine ordinale Variable ist eine Variable, die nur Aussagen über die Reihenfolge oder Größenordnung zulässt, z. B. Schulnoten, "ratings", Rangplätze. Wir nehmen an, dass große numerische Werte der positiven Ausprägung des bewerteten Merkmals entsprechen. Zwei ordinale Variablen sind (strikt) isoton (antiton), wenn sie in derselben Richtung (strikt) wachsen (abnehmen). So sind die folgenden Rangreihen isoton (oder [schwach] isoton, betont den Gegensatz zu strikt isoton): 1,2,3,4,5,6; 1,1,3,4,4,6;1,2,2,2,5,6. Eine Variable wächst isoton mit einer zweiten, wenn sie nicht abnimmt (gleich bleibt oder wächst), wenn die zweite wächst. Eine Variable X "dominiert stochastisch" eine zweite Y,  $X \succ Y$ , wenn ihre kumulative Verteilungsfunktion  $F(x) = P(X \le x)$  zur Gänze "unterhalb der von Y liegt",  $F(x) \le G(Y = x)$ 

(G Verteilungsfunktion von Y) für alle Werte x. Wenn zwei Schüler jeweils zwei Noten bekommen (in Latein und Mathematik), dann ist nicht derjenige mit der geringeren Summe oder dem niederen arithmetischen Mittelwert der bessere, sondern derjenige, der in beiden Fächern die kleinere Note hat, falls es einen solchen gibt (s. unten); wenn in einem Fach der eine Schüler besser ist und im anderen der andere, dann sind sie bezüglich ihrer Gesamtleistung nicht ordenbar (partielle Ordnung).

Wir definieren zunächst das messtheoretische Axiomensystem und stellen dann in Form von Theoremen seine Eigenschaften fest. Die Axiome beziehen sich auf eine Systemcharakteristik, eine Verteilungsfunktion, die Funktion von zwei Parametern, einem Personenparameter und einem Itemparameter ist, die die Wahrscheinlichkeit eines Zahlenwertes oder eines kleineren angibt. Diese Funktion mit einem festen Itemparameter, betrachtet als Funktion des Personenparameters, heißt Itemcharakteristik. Dann können wir gleich auch den Begriff der Personencharakteristik einführen. Das ist die Systemcharakteristik mit einem festen Personenparameter, betrachtet als Funktion eines variablen Itemparameters. Das Axiom W1 der schwachen Personenunabhängigkeit (weak subject independence; Irtel & Schmalhofer 1982, Irtel 1987) postuliert, dass sich die Personencharakteristiken nicht schneiden, d. h. dass die Personen bei jedem Item durch ihre Antwortverteilungsfunktionen isoton geordnet sind. Die Ordnung der Personen ist folglich unabhängig von den Items. Das Axiom W2 der schwachen Itemunabhängigkeit (weak item independence) postuliert nicht schneidende Itemcharakteristiken, dass also die Schwierigkeiten der Items für alle Personen isoton geordnet sind (Unabhängigkeit der Itemordnungen von den Personen). Das Axiom LI (local independence) der lokalen stochastischen Unabhängigkeit besagt, dass die Reaktionen einer Person auf mehrere Items stochastisch unabhängig sind. Die Theoreme beweisen die Existenz von zwei Ordinalskalen (Parametern, Personen bzw. Items) und, dass diese Skalen unabhängig vom Referenzsystem, also spezifisch objektiv sind.

*Definition:* Ein probabilistisches Paarvergleichssystem  $\langle A \times Q, P \rangle$ , wo P(t; a, x) eine Familie von Verteilungsfunktionen auf  $A \times Q$  ist, ist *schwach experimentell unabhängig geordnet*, wenn folgende Axiome gelten (Scheiblechner 1995):

- W1. Wenn P(t; a,x) > P(t; b, x) für ein x und ein t, dann P(t\*; a,y)  $\geq$  P(t\*; b, y) für alle y und alle t\* (schwache Personenunabhängigkeit).
- W2. Wenn P(t;a,x) > P(t;a,y) für ein a und ein t, dann  $P(t^*;b,x) \ge P(t^*;b,y)$  für alle b und alle  $t^*$  (schwache Itemunabhängigkeit).
- LI.  $P(t^{(k)}; a, x^{(k)}) = \prod_{x=1}^{k} p(t; a, x),$ wobei  $t^{(k)}$  und  $x^{(k)}$  Vektoren der Länge k sind (*lokale Unabhängigkeit*).

Der Testdatensatz heißt instrumentelles Paarvergleichssystem, weil Paarverglei-

che zwischen Personen bzw. Items durchgeführt werden mit den "Instrumenten" der Items bzw. Personen. Die Systemcharakteristik ist die Familie der Verteilungsfunktionen auf  $A \times Q$ , wobei A = (a,b,...) die Menge der Personen und Q = (x,y,...) die Menge der Items ist. Die beiden Axiome W1 und W2 definieren die Bi-Isotonie.

*Theorem 1:* Ein endliches System  $\langle A \times Q, P \rangle$  ist genau dann schwach instrumentell unabhängig geordnet, wenn auf A bzw. Q Ordinalskalen  $\phi_A$  bzw.  $\phi_Q$  existieren, so dass

1a.  $\varphi_{\Delta}(a) > \varphi_{\Delta}(b) \Leftrightarrow P(t; a, x) < P(t; b, x)$  für einige x und einige t und

1b. 
$$\varphi_0(x) > \varphi_0(y) \Leftrightarrow P(t; a, x) < P(t; a, y)$$
 für einige a und einige t.

P(t;a,x) < P(t;b,x) impliziert, dass a den kleineren Skalenwert hat, während P(t;a,x) = P(t;b,x) nichts impliziert. Wenn das endliche System als repräsentativ für entsprechende Populationen betrachtet wird, dann gilt die Aussage für Populationen. Wenn die Prämissen in W1 und W2 für  $\geq$  gelten, dann ist das System *strikt experimentell unabhängig geordnet* und die Sätze 1a und 1b gelten ebenso für  $\leq$  *und*  $\geq$  (strikt isoton).

Theorem 2: Unabhängigkeit bezüglich Subsystemen (Referenzsystemen).

Die beiden Ordinalskalen sind unabhängig von den Mengen A bzw. Q.

Sofern zwei Personenstichproben dieselben Personen a, b enthalten, ist die Ordnung ihrer Skalenwerte bis auf mögliche Bindung unabhängig von der Itemstichprobe (eine Itemmenge diskriminiert nicht bezüglich zweier Ausprägungsgrade). Analoges gilt für die Items. Die Skalen sind bezüglich der Messinstrumente spezifisch objektiv.

Theorem 3: Wenn das polytome instrumentelle Paarvergleichssystem schwach unabhängig geordnet ist, dann ergibt jede beliebige Dichotomisierung der instrumentellen Variablen ein schwach geordnetes dichotomes instrumentelles Paarvergleichssystem.

Es ist sogar möglich, jedes Item oder jede Vp an verschiedenen Schnittstellen (cutting scores) zu dichotomisieren, etwa um Niveauunterschiede auszugleichen.

*Theorem 4:* Die Rohwertverteilungen von Personen mit höheren Skalenwerten (Items mit größeren Parametern) dominieren stochastisch die Rohwertverteilungen von Personen mit niederen Skalenwerten.

Der Rohwert einer Person (Item) ist die Summe ihrer ordinal kodierten Itemantworten. Der Rohwert korreliert meist hoch positiv mit dem ordinalen Skalenwert. Eine Änderung der Itemauswahl kann zuvor diskriminierte Personen ununterscheidbar machen. Das Rohwertsystem ist ein Paarvergleichssystem, bei dem die Personen bzw. Items nach ihren Randsummen geordnet sind. Im dichotomen Fall ist auch das Rohwertsystem ein schwach geordnetes System, falls das "wahre" System ein schwach geordnetes instrumentelles System ist.

Ein ISOP ist eine echt probabilistische Mess-Struktur (keine sekundäre, aufgesetzte Fehlerverteilung) und kann daher nach den üblichen statistischen Kriterien getestet werden. Das ISOP ist eine nichtparametrische Verallgemeinerung vieler parametrischer probabilistischer Modelle (exponentielle Verteilungsfamilien, Poisson, Exponentialverteilung) und deterministischer ordinaler Modelle (Guttman, Mokken) und kann daher auch als (notwendiger) Test dieser verwendet werden.

Wenn man den Axiomen noch ein Kürzbarkeitsaxiom für Personen und Items hinzufügt, nämlich dass man einen Fähigkeitszuwachs der Personen durch einen Schwierigkeitszuwachs der Items ausgleichen kann, dann erhält man ein additiv verbundenes System, ADISOP (additives ISOP), mit Intervallskalen für die Personen und Items und einer additiven Verknüpfung der Skalenwerte. Wenn man den Axiomen ein instrumentelles Unabhängigkeitsaxiom hinzufügt, dass nämlich die Bewertung (Antwortskala, Scoring) der Antworten unabhängig von Personen und Items ist, dann erhält man ein vollständig additives System, CADISOP (vollständig additiv verknüpftes ISOP), mit Intervallskalen für die Personen, Items und Antwortvariablen, deren Skalenwerte additiv verknüpft sind (das eindimensionale mehrkategoriale Modell von Rasch ist ein parametrisches Beispiel für eine derartige Messstruktur). Die zusätzlichen Axiome lassen sich im Gegensatz zu den Axiomen von ISOP nicht direkt und unabhängig von den anderen prüfen, sondern nur durch Anpassung von ISOP, ADISOP und CA-DISOP, Berechnung der Likelihood der Daten unter den Modellen und durch deskriptive Likelihoodguotienten.

Die Prüfung der Modellgeltung von ISOP erfolgt durch die statistische Prüfung der Geltung der Axiome (Scheiblechner 2003). Der Goodman-Kruskal-Index (1954) G oder  $\gamma$  der ordinalen Assoziation prüft die (schwach) isotone Korrelation zweier ordinaler Variablen X und Y

$$G = \frac{N_s - N_d}{N_s + N_d} \text{ ,} \label{eq:gamma_def}$$

wobei  $N_s$  die Anzahl der Paare von Messwerten ist, die in X und Y gleichgerichtet streng geordnet (streng isoton) sind (similarly ordered), und  $N_d$  die Anzahl der Paare, die (streng) antiton geordnet sind (discordantly ordered).

G ist asymptotisch normalverteilt mit dem wahren Wert  $\gamma$  als Erwartungswert und bekannter Varianz. Daher kann man die (positive) Abweichung von 0 oder jedem anderen gegebenen Wert auf Signifikanz testen oder ein Konfidenzintervall um die Stichprobenschätzung errichten.

Das ISOP-Modell impliziert nichtnegative Korrelationen für alle Itempaare, alle Paare von Personen (über die Items) und alle Subsysteme des Referenzrahmens. Wenn man die Personen oder die Items nach mehreren (mehr als zwei) Variablen (in zwei partiellen Ordnungen) gleichzeitig ordnen kann (s. unten), dann müssen diese partiellen Ordnungen bei ISOP ebenfalls nichtnegativ korreliert sein. So kann

man beispielsweise prüfen, ob die Ordnungen der Personen durch zwei verschiedene Itemmengen übereinstimmen (spezifisch objektiv sind). Man braucht nur die gleichgerichteten und die entgegengesetzt gerichteten Ordnungen auszuzählen Dies ist eine simultane Korrelation von  $r \ge 1$  unabhängigen X-Variablen mit  $s \ge 1$  abhängigen Y-Variablen.

Diese Testgrößen wurden auf die Konstruktion eines Tests für den Kohärenzsinn angewandt (MR-SOC, Marburger Test des sense of coherence; Scheiblechner & Lutz 2009). Der MR-SOC misst die Verbundenheit einer Person mit ihrem alltäglichen sozialen Umfeld, die Kohärenz. Diese ist hoch, wenn eine Person das Gefühl hat, an sie gerichtete soziale Anforderungen gut vorhersehen zu können; in der Lage zu sein, diese Anforderungen erfüllen zu können und dass es lohnend ist, sich den Anforderungen zu stellen.

Item-Nr.		Wortlaut der Items des MR-SOC	Position	Korrelation
neg	pos		Darbietung	G N P
1		Das hat ja doch keinen Zweck.	1	.67 . <b>74</b> .40
2		Mir gleitet alles aus den Händen.	2	.62 . <b>71</b> .36
	11	Wie ist das Leben schön.	3	.63 .38 . <b>69</b>
3		Ich weiß einfach nicht, was ich will.	4	.50 . <b>58</b> .29
4		Alles ist so unbefriedigend.	5	.63 . <b>71</b> .38
5		Ich weiß nicht, was ich davon halten soll.	6	.49 . <b>65</b> .22
6		Ach, ist doch egal.	7	.47 . <b>57</b> .26
	12	Mensch, wie toll.	8	.61 .35 . <b>71</b>
	13	Ich weiß, wo's lang geht.	9	.58 .33 . <b>71</b>
	14	Ich bin schon mit mehr fertig geworden.	10	.44 .22 . <b>55</b>
7		Ich versteh das einfach nicht.	11	.54 . <b>63</b> .30
	15	Das macht Spaß.	12	.66 .40 . <b>75</b>
	16	Ich laß' mich davon nicht unterkriegen.	13	.55 .32 . <b>64</b>
8		Ich bin ganz durcheinander.	14	.57 . <b>70</b> .30
	17	Das schaff' ich schon.	15	.67 .43 . <b>73</b>
9		Ich kann das nicht.	16	.56 . <b>66</b> .31
	18	Ich weiß, wie es weitergeht.	17	.58 .30 . <b>74</b>
10		Dagegen bin ich machtlos.	18	.52 . <b>66</b> .26
	19	Mir wird schon eine Lösung einfallen.	19	.56 .30 . <b>69</b>
	20	Es ist zwar anstrengend, aber es lohnt sich.	20	.53 .26 . <b>69</b>

Tabelle 1: MR-SOC-Fragebogen (Lutz 2001)

Anmerkungen: Die Items sind Antworten auf die Formulierung "Ich sage zu mir in dieser oder einer ähnlichen Art:", die mit 1 (nie), 2 (selten), 3 (öfter), 4 (häufig) für die letzten zwei Wochen zu beantworten sind. Für die negativen Items wurde die Antwortnummerierung umgedreht.

Position: tatsächliche Darbietungsreihenfolge, Item-Nr.: im Text verwendete Nummerierung, pos/neg: Zuweisung zur positiv bzw. negativ formulierten Testhälfte des Persönlichkeitsmerkmals.

Tabelle 1 enthält die Items des MR-SOC und ihre isotonen Korrelationen mit den partiellen Ordnungen der Personen nach dem Gesamttest G und den Untertests P der positiv formulierten Items bzw. N der negativ formulierten Items. Alle Items werden ohne Kenntnis des Wortlauts nach der Höhe der Korrelationen korrekt den Untertests zugewiesen.

Die gewogenen mittleren Korrelationen der isotonen Ordnungen der Personen nach den einzelnen Items (klassisch testtheoretisch etwa Itemreliabilitäten oder Itemvaliditäten) ergeben folgende Prüfgrößen für die Geltung des Axioms W1 im Gesamttest bzw. in den Untertests:

Test	Prüfgröße	Standardabweichung
N	0.544	0.111
Р	0.643	0.176
G	0.410	0.108

Alle weichen signifikant positiv von 0 ab und die Wahrscheinlichkeiten zufällig so hoher Werte liegen unter dem Promillebereich. Es ist zu beachten, dass sich die Prüfgrößen auf Tests der Länge 1 beziehen.

Analog lauten die Prüfgrößen für die Geltung des Axioms W2:

Test	Prüfgröße	Standardabweichung
N	0.090	0.047
P	0.216	0.073
G	0.149	0.067

Die geringe Größe der Prüfgrößen für W2 ist Ausdruck der geringen Schwierigkeitsunterschiede der Items. Selbst für N ist die Wahrscheinlichkeit, dass W2 nicht gilt (= dass die Prüfgröße 0 oder negativ ist), nur 2.8% (z = -0.09/0.047).

# 2 Mehraspekt-Modelle: d-dimensionale Testdaten

In einigen Sportarten werden die dargebotenen Leistungen nach mehreren Aspekten gleichzeitig bewertet (Skispringen: Weite, Haltung; Eiskunstlauf: athletischer Wert, künstlerischer Eindruck) und schließlich wird eine kombinierte gemeinsame Rangreihe der Sportler erstellt. Bei intellektuellen Leistungen (und erst recht in ökologischen Alltagssituationen) ist nicht nur die Angemessenheit einer Problemlösung, sondern auch die Geschwindigkeit (und mögliche weitere Aspekte) entscheidend (Speed-Power-Problem).

*Definition:* (Mehrdimensionale) koordinatenweise partielle Ordnung (Scheiblechner 1999). Eine Menge  $X = R^d$  mit Elementen  $x = (x_1, ..., x_i, ..., x_d)$ , wobei  $x_i$  reelle Zahlen sind, ist eine (mehrdimensionale) koordinatenweise partielle Ordnung, wenn sie mit der Ordnung

$$x \succ \approx y$$
 gdw  $x_1 \geq y_1, x_2 \geq y_2, ..., x_d \geq y_d$ 

ausgestattet ist.

*Definition:* d-dimensionale stochastische Dominanz von Variablen oder Verteilungen (concordance ordering)

$$X \succ \approx Y \text{ oder } F_x \succ \approx F_y$$

gdw  $F_X(x_1,...,x_i,...,x_d) \leq F_Y(y_1,...,y_i,...,y_d)$  für alle  $(x_1,...,x_i,...,x_d) \in \mathbb{R}^d$ , wobei  $F_X(x_1,...,x_2,...,x_d)$  die kumulative Verteilungsfunktion (c.d.f.) von X in d Dimensionen ist. Wenn die Kleiner-Relation in mindestens einer Dimension strikt ist, heißt die stochastische Dominanz strikt.

Im Axiomemsystem für das eindimensionale polytone ISOP-Modell braucht man nur die eindimensionalen c.d.f.s P(t; a, x) durch mehrdimensionale kumulative Verteilungsfunktionen zu ersetzen:

$$F_X$$
 (x; a, x)X, x...d-dimensionale Vektoren.

Die Beweise und Theoreme sind analog.

Eine Logistikfirma suche Kraftfahrer für Ferntransporte. Zwei Eigenschaften seien maßgeblich: A Fahrsicherheit inklusive Disziplin (Psychomotorik, Alkohol, Pausen ...), B Logistik (Routenplanung, internationale Verkehrs- und Transportvorschriften ...). Es sollen nur die besten 20% der Bewerber eingestellt werden.

Für den zweiten Teil eines Fortbildungskurses sollen die schlechtesten 50% der Teilnehmer des ersten Teils ausgeschlossen werden. Es liegen zwei Leistungen vor: A schriftliche Prüfungsleistung, B mündliche Prüfungsleistung.

Punkte A\B	B: 1	2	3	4	5	Summe
A: 1	4	4	2	3	2	15
2	2	4	8	3	2	19
3	4	7	10	7	4	32
4	2	3	8	4	4	21
5	2	3	2	2	4	13
Summe	14	21	30	19	16	100

Tabelle 2: Häufigkeiten von zwei Prüfungsleistungen oder Qualifikationen

In einem eindimensionalen Test mit mehreren Items ist der (modifizierte) Perzentilscore (P-Score) eine ordinale Maximum-Likelihood-Schätzung der Rangposition einer Person oder eines Items:

$$r(x;P,Q) = \frac{n^+ n}{n^+ + n}$$

wobei n<sup>+</sup> die Anzahl kleinerer ("übertroffener") Konkurrenten und n<sup>-</sup> die Anzahl größerer ("überlegener") Konkurrenten ist.

Der P-Score benutzt die Häufigkeiten von Maßzahlen zu einer vergleichenden Messung, wie sie in psychologischen Tests üblich ist. Er vermeidet Vergleiche unterschiedlicher Dimensionen, die Verwendung von subjektiven Wahrscheinlichkeiten und von Lotterien, wie es bei der kardinalen Nutzenmessung notwendig ist.

Bei Mehraspektmerkmalen ist der Score gleich definiert, wobei jedoch die mehrdimensionale Größer-kleiner-Relation zu verwenden ist.

Wir berechnen nun die Perzentilscores der 2-Aspekt-Leistungen von Tabelle 2.

$\sum_{y < x} Y = n^+$ A:	B:1	2	3	4	5	$\sum_{x < y} Y = n^{-}$ A:	B: 1	2	3	4	5
1	0	4	8	10	13	1	96	82	63	32	14
2	4	10	16	27	32	2	83	71	50	27	12
3	6	18	35	51	62	3	62	51	35	18	8
4	10	27	50	71	83	4	32	27	16	10	4
5	12	32	63	82	96	5	11	8	6	4	0

Tabelle 3: Berechnung der Anzahl übertroffener und unterlegener Fälle

Modifizierter Perzentilscore	B: 1	2	3	4	5
A: 1	-1	92	822	524	087
2	908	753	515	0	.455
3	824	478	0	478	.771
4	524	0	.515	.753	.908
5	0.043	.6	.826	.907	1

Tabelle 4: Modifizierte Perzentilscores der Tabelle 2

Wir suchen die 20% Besten und die 50% Schlechtesten (vgl. Tab. 5 & 6). Bis zur Zelle (1,5) mit Score -.037 (vgl. Tab. 6) haben wir nur 38%, dann folgen 16 % mit Score 0, deren Einschluss 54% ergäbe. Wir bestimmen per Los die unglücklichen fehlenden 12%. Wir sehen, dass Personen mit der gleichen Durchschnittsnote verschiedene Scores bekommen können. Für das Testmanual eines Mehraspekttests ist eine lexikografische Liste mit den Maßzahlvektoren und den zugeordneten Scores zur Verfügung zu stellen. Wir sehen auch, dass die Ordnung

der Scores häufig nicht stark von der Ordnung der Mittelwerte abweicht, was die Beliebtheit der letzteren erklärt, obwohl Rohwert und Mittelwert zahlreiche unplausible und ungerechtfertigte Annahmen machen: dass die Abstände zwischen je zwei benachbarten Rangzahlen alle gleich sind, dass alle Aspekte die gleiche Intervallskala aufweisen.

Zelle (Vektor)	Häufigkeit	Perzentilscore	kumulative Häufigkeit
(5,5)	4	1	4
(4,5)	4	.908	8
(5,4)	2	.907	10
(5,3)	2	.826	12
(3,5)	4	.771	16
(4,4)	4	.753	20

Tabelle 5: Die 20% Besten

Zelle (Vektor)	Häufigkeit	Perzentilscore	kumulative Häufigkeit
(1,1)	4	-1	4
(1,2)	4	92	8
(2,1)	2	908	10
(3,1)	4	824	14
(1,3)	2	822	16
(1,4)	3	524	19
(4,1)	2	524	21
(2,3)	8	515	29
(3,2)	7	478	36
(1,5)	2	037	38
(2,4)	3	0	
(3,3)	10 } 16	0	
(4,2)	3	0	54

Tabelle 6: Die 50% Schlechtesten

Der SSS-Test (small group social satisfaction test; Lutz 2001) klassifiziert die täglichen Face-to-face-Interaktionspartner in sechs Gruppen: Nachbarn und Bekannte, (entfernte) Verwandte, Freunde, Arbeitskollegen, Familienmitglieder, Lebenspartner (falls vorhanden). Für jede Gruppe werden drei Aspekte abgefragt: Häufigkeit positiver Kontakte, Häufigkeit negativer Kontakte, Häufigkeit neutraler Kontakte. Die Häufigkeiten werden in fünf Abstufungen angegeben: nie, selten, manchmal, häufig, sehr häufig. Der Test wurde bei 1.232 Personen angewandt. Die Korrelationen zwischen den Items sind durch den Goodman-Kruskal-Index der

Assoziation zwischen den partiellen Ordnungen oder Perzentilskores der beiden Items definiert. Sie sind hoch signifikant positiv zwischen allen Items (Minimum 0.108 zwischen Kollegen und Lebenspartnern, Maximum 0.709 zwischen Bekannten und Freunden, typisch über 0.4). Die Item-Rest-Korrelationen sind die mehrdimensionalen Korrelationen der partiellen Ordnungen durch ein Item und der Ordnung durch die übrigen Items. Sie betragen zwischen 0.458 (Lebenspartner) und 0.812 (Familienmitglieder) und sind hoch signifikant positiv. Die Kriterien für die Gültigkeit des Axioms W1 für die Items und den Gesamttest betragen zwischen 0.458 (Lebenspartner) und 0.812 (Familienmitglieder) und sind hoch signifikant positiv. Das Kriterium für W1 entspricht in etwa der klassischen Trennschärfe. Das Testkriterium für W2 beträgt 0.224 (normalverteilt, Standardabweichung 0.0008; geringe Schwierigkeitsunterschiede der Items). Der SSS-Test entspricht sehr gut den Forderungen an einen eindimensionalen Test (Familienmitglieder und Lebenspartner gehören in den Bereich der privaten sozialen Zufriedenheit und können von den anderen Gruppen getrennt werden, die in den Bereich der öffentlichen sozialen Zufriedenheit gehören).

Der Erfolg des Verhaltens hängt unter ökologischen Bedingungen meist von vielen Aspekten gleichzeitig ab. Psychologische Tests sollten sich diesem Prinzip endlich annähern. Das messtheoretische d-ISOP-Modell erlaubt eine Modellgültigkeitsprüfung ohne numerische Modellanpassung. Die beobachteten d-dimensionalen Reaktionen können in einer eindimensionalen latenten Dimension angeordnet werden. Der modifizierte Perzentilskore erlaubt für jede gegebene Ausgangslage eine Aussage, durch welche Verbesserung in welcher Beobachtungsdimension das Gesamtergebnis am meisten gesteigert werden kann.

#### 3 Multiattributive Entscheidungstheorie: Wer sind die Besten?

36 Athleten bewarben sich bei den Olympischen Spielen in Peking 2008 um die Medaillen im Zehnkampf. 26 beendeten den Wettkampf. Es soll der perfekteste Leichtathlet gefunden werden, der die Einzeldisziplinen am vollständigsten absolviert hat. Jedoch gibt es keinen, der überall der Beste ist. Wie kann man aber eine Hundertstelsekunde beim 100-m-Lauf mit einem Zentimeter im Hochsprung vergleichen? Selbst innerhalb der Disziplinen, die alle durch Zeiten oder alle in Zentimetern gemessen werden, ist der Vergleich unklar: Wie viele Hundertstel im 100-m-Lauf entsprechen einer Sekunde im 1.500-m-Lauf? Ist der Abstand zwischen 9.6 und 9.7 Sekunden im 100-m-Lauf gleich groß wie der zwischen 10.6 und 10.7 Sekunden? Die physikalischen Maßeinheiten entsprechen nicht vergleichbaren Leistungsunterschieden.

Ein Konkurrent ist besser als ein anderer, wenn er in allen Disziplinen gleich gut wie dieser und in mindestens einer besser ist (d-dimensionale Dominanz). Viele Paare von Athleten können so nicht geordnet werden, es sollen aber auch nicht Abstände in unterschiedlichen Disziplinen miteinander verglichen werden. In der offiziellen Wertung werden die Leistungen durch willkürliche Transformationen

(für welche man auch Informationen über die Disziplin benötigt) in Punkte umgerechnet, die addiert werden.

Weiters ist auch unklar, welche Disziplinen zum Zehnkampf gehören sollen. Weshalb ist kein Marathon dabei? Weshalb kein Gewichtheben?

Das polytome eindimensionale ISOP-Modell misst eine übliche psychologische Dimension. Im d-ISOP werden verschiedene Dimensionen (Aspekte) beobachtet, aber auf eine einzige zugrunde liegende Dimension zurückgeführt. Der Zehnkampf soll nicht nur eine einzige Leistung messen, aber auch keine willkürliche Zusammenstellung unvereinbarer Anforderungen sein. Gesucht ist die beste (gerechteste) eindimensionale Näherung an ein Bündel verwandter Dimensionen.

Wir untersuchen die isotonen Zusammenhänge der Disziplinen. Die Disziplinen sollen nicht negativ miteinander korrelieren, aber auch nicht nur hoch positiv, denn der Zehnkampf soll vielfältig sein. Die Goodman-Kruskal-Indizes der Assoziation liegen zwischen 0.448 (100 m und 110 m Hürden) und -0.212 (1.500 m und Diskus). 11 der 45 Korrelationen sind negativ, keine signifikant, 12 der 34 positiven Korrelationen sind signifikant. Der 1.500-m-Lauf weicht am stärksten von den übrigen Disziplinen ab, hat aber immer noch eine 0.182-Assoziation zwischen der Disziplinrangordnung und der partiellen Ordnung der Athleten durch die anderen Disziplinen. Diese können in drei Gruppen eingeteilt werden:

- Laufen (100 m, 400 m, 110 m Hürden)
- Springen (Weitsprung, Hochsprung, Stabhochsprung)
- Werfen (Kugelstoßen, Diskuswurf, Speerwurf)

Diesen Gruppen entsprechen auch grob einige Sportlertypen. Für den 1.500-m-Lauf stellt sich die Frage, ob man ihn eliminieren sollte, um zu einem Neunkampf zu kommen, oder ob man ihn gerechterweise ebenfalls auf drei Disziplinen (Ausdauerleistung) aufwerten sollte (z. B. ähnlich dem Triathlon: Schwimmen, Radfahren, Marathon).

Der Rang9 (Punktsumme nach 9 Bewerben, ohne 1.500 m) unterscheidet sich maximal um 5 Positionen (Athleten 11, 13) vom endgültigen Rangplatz, der Score10, der Perzentilscore nach 10 Einzeldisziplinen um 6 Positionen (Athlet 6) (vgl. Tab. 7). Bei 10 Attributen gibt es nur wenige paarweise Dominanzrelationen, aber der Score10, der die Disziplinen als 10 wiederholte Schätzungen *einer* Dimension (wie einen eindimensionalen Test mit 10 Items) interpretiert, hat keinerlei Schwierigkeiten. Der ScoreLSWA fasst zunächst die Disziplinengruppen L, S, W, A zu jeweils einem Perzentilscore zusammen und dann die 4 Scores zu einem Endresultat. Dieser Score erzeugt eine Überbewertung von A, 1.500 m, weil es gleich bewertet wird wie L, S, W, obwohl es nur eine Disziplin enthält. Der Score führt auch zu einer Umkehr der ersten drei Positionen (Medaillengewinner), die möglicherweise die letzte Disziplin nicht mehr gleich gewichtet haben (besonders der Gesamtsieger, Rang 23 in A, ein Anti-A-Typ). Der ScoreLSWA unterscheidet sich maximal um 6 Positionen (Athlet 6, ein Springer, Platz 2 in S, Platz 19 in L) vom offiziellen Endresultat und um 5 Positionen vom Score10 (Athlet 13). Die absolute Streubreite

von Score bringt auch zum Ausdruck, wie gut sich die jeweiligen Attribute zu einer einzigen Dimension zusammenfassen lassen. Die Streubreiten für die Gruppen sind 1.803 für S, 1.787 für W und 1.547 für L (relativ abweichend 400 m) und für die Gesamtscores 1.38 für ScoreLSWA und 1.302 für Score10.

Athlet	Rang9	Score10	ScoreLSWA	ScoreL	ScoreS	ScoreW
1	1	-0.638 (1)	-0.520 (3)	-0.760	-0.706	-0.920
2	2	-0.467 (2)	-0.560 (2)	-0.595	-0.857	0.173
3	4	-0.443 (3)	-0.600 (1)	-0.627	-0.362	-0.227
4	3	-0.271 (5)	-0.180 (8)	0.253	-0.694	-0.707
5	7	-0.182 (9)	-0.400 (5)	0.081	-0.118	-0.333
6	5	-0.095 (12)	-0.080 (12)	0.425	-0.743	-0.173
7	10	-0.336 (4)	-0.455 (4)	-0.676	0.000	-0.147
8	9	-0.135 (10)	-0.180 (8)	0.014	-0.667	0.333
9	12	-0.262 (6)	-0.360 (6)	-0.680	0.217	-0.200
10	14	-0.207 (7)	-0.313 (7)	-0.270	-0.529	0.413
11	6	-0.062 (14)	-0.100 (11)	0.351	-0.275	-0.387
12	11	-0.128 (11)	-0.180 (10)	0.547	-0.493	-0.467
13	8	-0.190 (8)	-0.082 (13)	-0.547	-0.164	-0.040
14	15	0.099 (16)	0.071 (15)	0.493	-0.045	-0.040
15	13	-0.083 (13)	0.061 (14)	-0.547	0.194	-0.093
16	16	0.176 (19)	0.152 (16)	0.467	0.343	-0.147
17	17	0.276 (11)	0.300 (18)	0.760	0.471	-0.360
18	19	0.143 (17)	0.180 (17)	0.108	0.352	0.067
19	20	0.159 (18)	0.260 (20)	-0.333	0.429	0.440
20	18	0.045 (15)	0.232 (19)	-0.573	0.194	0.227
21	22	0.240 (20)	0.280 (22)	-0.227	0.200	0.787
22	21	0.320 (22)	0.360 (24)	0.200	0.889	-0.280
23	23	0.412 (24)	0.340 (23)	0.622	-0.014	0.867
24	24	0.664 (26)	0.780 (26)	0.787	0.565	0.680
25	25	0.363 (23)	0.253 (21)	0.054	0.946	0.413
26	26	0.577 (25)	0.740 (25)	0.680	0.803	0.120

Tabelle 7: Die Ränge der Athleten und die Perzentilscores

Die Athleten werden nach ihrer offiziellen endgültigen Position nummeriert. Rang9 ist ihre Position nach 9 Disziplinen (ohne 1.500 m). Je niederer der Rang und je negativer der Perzentilscore, desto besser die Leistung. Score10 (Rang in Klammern) ist der Perzentilscore, basierend auf der partiellen Ordnung der 10 Einzelleistungen. Score LSWA ist der Perzentilscore beruhend auf den Perzentilscores der Gruppen L, S, W, A (Ausdauer, beruht nur auf 1.500 m). (Maxima und Minima kursiv).

Obwohl die Summe der transformierten Punkte für die Endwertung der Athleten herangezogen wird, korrelieren die P-Scores, die nur die Ranginformationen der Athleten im aktuellen Wettkampf benutzen, ähnlich hoch mit dieser (Score S 0.848, Punktsumme S 0.791). Es wäre kein Problem, die Ergebnisse zurückliegender Wettbewerbe und einer größeren Athletengruppe einzubeziehen. Die Endposition korreliert zu 0.964 mit Rang9, zu 0.934 mit Score10 und zu 0.928 mit dem ScoreLSWA. Die Scores der Gruppen korrelieren niedriger miteinander als die entsprechenden Punktsummen, was wieder die P-Scores und die vorgenommene Einteilung der Disziplinen unterstützt.

# Ein marginaler Ansatz zur Schätzung von Item-Response-Modellen mit zufälligen individuellen Effekten

Martin Spiess

#### 1 Zur Schätzung von Item-Response-Modellen

Die Psychologie ist eine empirische Wissenschaft, in der die Verknüpfung von Theorie und Empirie ein zentraler Baustein ist. Diese Verknüpfung liefern die psychologischen Methoden, die selbst wissenschaftlichen Kriterien genügen müssen. Mit Kempf (2003a) sind wissenschaftliche Methoden:

"... Instrumente des systematischen Zweifels an der Wahrheit von Behauptungen. Sie zwingen uns, für das, was wir zu wissen meinen, Gründe vorzubringen. Und sie nehmen uns zugleich in die Pflicht, die Grenzen unseres Wissens zu explizieren, d. h. insbesondere auch: seine impliziten Voraussetzungen transparent zu machen." (S. 27)

Dies gilt in gleichem Maße für die statistischen Methoden, die in der Psychologie eine weite Verbreitung gefunden haben. D. h. es gilt auch für die statistischen Methoden, dass deren Wissenschaftlichkeit nicht an der unreflektierten Abarbeitung bestimmter, vorgegebener Schritte festzumachen ist (vgl. ebd.), sondern dass auch diese Verknüpfung oder Argumentationskette ausgehend von Annahmen hin zu Aussagen mit wissenschaftlichen Kriterien zu beurteilen ist. Ein solches, wenn auch nicht das wichtigste Kriterium ist die Einfachheit, die im Methodenkontext folgendermaßen interpretiert werden kann: Wenn es zwei Methoden gibt, die Schlüsse gleicher Qualität erlauben, dann ist die mit weniger Aufwand verbundene Methode vorzuziehen. Es versteht sich von selbst, dass eine Methode A dann einer Methode B vorzuziehen ist, wenn A Aussagen höherer Qualität erlaubt. Dabei ist der Begriff Qualität näher zu definieren: Er meint hier die statistischen Eigenschaften der Schätzer, aber auch die Robustheit der Verfahren gegenüber Annahmeverletzungen sowie die technische Robustheit bei der Berechnung der Schätzer. Gegenstand dieses Beitrages ist der Vergleich zweier Schätzverfahren im Kontext probabilistischer Testmodelle entsprechend diesen Kriterien.

Die vorliegende Arbeit beschäftigt sich mit der Schätzung der Itemschwierigkeitsparameter insbesondere in binären Item-Response-Modellen. Die übliche Praxis besteht darin, die logarithmierte Likelihoodfunktion zu formulieren und den Maximum-Likelihood-(ML-)Schätzer, d. h. den plausibelsten Schätzwert für die beobachteten Daten zu berechnen. Sind alle notwendigen Annahmen erfüllt, dann besitzt der ML-Schätzer günstige Eigenschaften: Er ist konsistent, d. h. mit größeren

Stichproben steigt die Wahrscheinlichkeit, dass die Schätzwerte nahe bei dem wahren Wert liegen, er ist asymptotisch effizient, also für große Stichproben recht präzise, und asymptotisch normal verteilt.

In vielen Fällen ist die Likelihoodfunktion allerdings nicht mehr analytisch berechenbar. Stattdessen sind numerische Methoden notwendig, die es erlauben, die zu maximierende logarithmierte Likelihoodfunktion bzw. deren Ableitungen näherungsweise zu berechnen. Die günstigen Eigenschaften der ML-Schätzer können aber nur in Anspruch genommen werden, wenn alle notwendigen Annahmen erfüllt sind und die numerische Annäherung hinreichend genau ist. Im Allgemeinen ist nicht bekannt, wie gut die Annäherung im Einzelfall ist. Eine Strategie kann dann darin bestehen, die Schätzer und deren Standardfehler mit einer steigenden Präzision zu berechnen. Ändern sich diese Werte nicht mehr oder fast nicht mehr, dann kann dies ein Hinweis auf eine hinreichend gute Annäherung sein. Diese Vorgehensweise ist allerdings rechenzeitaufwendig und fehleranfällig.

Im Folgenden wird eine Schätzmethode vorgestellt und evaluiert, die nicht auf der ML-Methode basiert. Sie erlaubt die einfache, analytische Schätzung der Itemschwierigkeitsparameter in speziellen, aber interessanten Modellen. Es lässt sich zeigen, dass die Eigenschaften dieser Schätzer denjenigen der ML-Schätzer gleichkommen. Ihre Berechnung ist einfach, schnell und robust. Darüber hinaus sind sie nicht, wie die numerisch gewonnenen ML-Schätzer, anfällig für Näherungsfehler. Sie stellen so eine attraktivere Möglichkeit dar, die Itemschwierigkeitsparameter zu schätzen.

Bei der Modellierung des Zustandekommens von Testleistungen wird üblicherweise von einem endlich oder unendlich großen Pool von Testaufgaben ausgegangen. Je nach Paradigma wird jeder Person eine Zufallsstichprobe aus diesem Itempool vorgelegt (Item-Sampling-Paradigma) oder alle Personen lösen dieselbe Menge an Items (Fixed-Test-Paradigma) (vgl. Kempf 1992b). Letzteres ist der in diesem Papier gewählte Ansatz. Weiterhin werden die in der Stichprobe beobachteten Personen bzw. genauer die durch die Stichprobenziehung realisierten Ausprägungen auf der/den entsprechenden Fähigkeitsdimension/en als Realisation einer Zufallsvariablen aufgefasst. Bei diesen hier betrachteten Item-Response-Modellen mit zufälligen individuellen Effekten wird üblicherweise angenommen, dass deren Verteilung bis auf bestimmte Parameter bekannt ist (z. B. Andersen & Madsen 1977; Bock & Aitkin 1981; Bock & Lieberman 1970). Die Schwierigkeitsparameter der Items (Itemparameter) werden im Allgemeinen gemeinsam mit diesen Parametern über die Maximum-Likelihood-(ML-)Methode geschätzt. In Übereinstimmung mit dem Rasch-Modell ist ein häufig verwendetes Modell für die bedingte Lösungswahrscheinlichkeit eines Items, gegeben die unbeobachtete Fähigkeit, die logistische Funktion. Eine andere gängige Funktion ist die kumulative Normalverteilungsfunktion (z. B. Bock & Lieberman 1970).

Allerdings ist die ML-Schätzung dieser Modelle nicht unproblematisch, denn die individuellen Effekte müssen bei der marginalen ML-(MML-)Methode aus der bedingten Verteilung der Responsevariablen "herausintegriert" werden. Da das entspre-

110 Martin Spiess

chende Integral nicht mehr analytisch berechnet werden kann, kommen meist numerische Methoden zur Anwendung. Wird etwa eine eindimensionale latente (zufällige) Variable als normal verteilt angenommen, dann ist die approximative Berechnung der Integrale mittels der Gauss-Hermite-Quadratur möglich (z. B. Bock & Lieberman 1970). Die Anzahl an Stützstellen, die den Rechenaufwand bestimmen und die für eine möglichst genaue Berechnung der Integrale und damit für eine konsistente Schätzung sowohl der interessierenden Parameter als auch der Varianzen der Parameterschätzer benötigt wird (Butler 1985), hängt unter anderem von der Varianz der unbeobachteten Effekte ab und kann bei geringen Stichprobenumfängen und einer hohen Korrelation der jeweils an derselben Person erhobenen Beobachtungen sehr groß werden (z. B. Spiess & Hamerle 2000). Die Schätzung von Modellen unter Annahme mehrdimensionaler latenter Variablen ist mit dieser Technik im Allgemeinen nicht mehr praktikabel (Bock & Aitkin 1981; Spiess 1995). Zwar ist eine Schätzung mit Hilfe einer adaptiven numerischen Methode möglich, sie bleibt aber nach wie vor zeitaufwendig und ist auch theoretisch nicht unproblematisch (siehe die Diskussion in Schilling & Bock 2005). Schließlich ermöglicht diese Vorgehensweise keine einfache Generalisierung auf andere Verteilungsmodelle.

Um solche Probleme vermeiden zu können, bieten sich andere, nicht likelihood-basierte Schätzverfahren an. Im Allgemeinen besteht der Preis für flexiblere und robustere Verfahren allerdings in einem (asymptotischen) Effizienzverlust. Werden zur Schätzung verallgemeinerte Schätzgleichungen (generalized estimating equations, GEE; Liang & Zeger 1986) verwendet, dann lässt sich zeigen, dass dies nicht sein muss. Darüber hinaus können die entsprechenden GEE-Schätzer unter bestimmten Bedingungen analytisch, ohne numerische Integrationsverfahren berechnet werden.

In Kapitel 2 wird zunächst die betrachtete Klasse von Item-Response-(IRT-)Modellen beschrieben. Im Anschluss daran (Kapitel 3) wird der GEE-Schätzer allgemein und speziell für die hier betrachtete Modellklasse eingeführt. Mögliche Erweiterungen des Schätzansatzes werden in Kapitel 4 angesprochen. In einer Simulationsstudie (Kapitel 5) werden der GEE-Schätzer und der ML-Schätzer basierend auf einer adaptiven Gauss-Hermite-Quadratur miteinander verglichen. In Kapitel 6 werden Vorteile, aber auch Grenzen des GEE-Schätzers im Hinblick auf die Schätzung der IRT-Modelle abschließend diskutiert.

### 2 Das IRT-Modell

Betrachtet wird in diesem Kapitel ein binäres Modell. Wie in Kapitel 3 gezeigt wird, können die Ergebnisse der folgenden Abschnitte aber ohne Weiteres auf Modelle mit stetigen oder diskret metrischen Antworten übertragen werden. Ordinal oder nominal skalierte Antworten sollen nicht betrachtet werden.

Ausgangspunkt sind i = 1,...,n Personen, die k = 1,...,K Items bearbeiten. Die beobachtete binäre Antwort von Person i auf das k-te Item wird mit  $y_{ik}$  bezeichnet, wobei  $y_{ik}=1$  eine korrekte Lösung und  $y_{ik}=0$  eine falsche Lösung des Items angibt. Bei der Beantwortung des k-ten Items durch die i-te Person spielt lediglich die Schwierigkeit des k-ten Items,  $\beta_{k}$ , sowie die "Fähigkeit" der i-ten Person, das k-te Item zu lösen,  $\pi_{ik}$ , eine Rolle. Die Variable  $\pi_{ik}$  kann, wie im multidimensionalen Faktorenmodell, als Funktion von Ladungen und latenten Variablen aufgefasst werden.

Den üblichen Annahmen folgend (z. B. Schilling & Bock 2005), sei  $\pi_{ik}$  eine über die Personen, aber nicht über die Items unabhängig  $N(0,\sigma_k^2)$ -verteilte Zufallsvariable und  $\epsilon_{ik}$  eine über die Personen und Items und von allen  $\pi_{ik}$  unabhängig  $N(0,1-\sigma_k^2)$ -verteilte Zufallsvariable. Die Summe  $\nu_{ik}=\pi_{ik}+\epsilon_{ik}$  ist damit N(0,1)-verteilt und über die Personen unabhängig, nicht aber über Items. Person i löst das k-te Item nicht, wenn  $\nu_{ik} \leq \beta_k$  bzw.  $\epsilon_{ik} \leq \beta_k - \pi_{ik}$ .

Mit der zusätzlichen Annahme:  $\pi_{ik}$  und  $\epsilon_{ik}$  sind unabhängig von allen  $\beta_k$  (k = 1,...,K), ist

$$Pr(y_{ik} = 0 \mid \beta_k, \pi_{ik}, \sigma_k) = \Phi((\beta_k - \pi_{ik}) / \sqrt{1 - \sigma_k^2}), \tag{1}$$

mit  $\Phi(\cdot)$  der Verteilungsfunktion der Standardnormalverteilung. Dabei sind lediglich die Parameter  $\beta_k^* = \beta_k/\sqrt{1-\sigma_k^2}$  identifizierbar. Wegen der Symmetrieeigenschaft der Normalverteilung ist, mit  $\pi_{ik}^* = \pi_{ik}/\sqrt{1-\sigma_k^2}$ ,

$$Pr(y_{ik} = 1 \mid \beta_{k}^{*}, \pi_{ik}^{*}) = \Phi(-\beta_{k}^{*} + \pi_{ik}^{*}).$$
 (2)

Für eine GEE-Schätzung der Itemparameter werden allerdings nicht diese Wahrscheinlichkeiten, sondern es wird

$$E(y_{ik}) = Pr(y_{ik} = 1 \mid \beta_k) = \Phi(-\beta_k)$$
(3)

benötigt, die (marginale) Wahrscheinlichkeit, das k-te Item zu lösen.

In Matrixschreibweise wird der Antwortvektor der i-ten Person zu  $y_i = (y_{i1},...,y_{iK})'$  und für die gesamte Stichprobe zu  $y = (y'_{1},...,y'_{n})'$ . Entsprechend für die anderen Komponenten des Modells:

$$\begin{split} \pi_i &= (\pi_{i_1}, \dots, \pi_{i_K})' \qquad \text{ und } \qquad \pi = (\pi_{i_1}', \dots, \pi_{i_n}')', \\ \epsilon_i &= (\epsilon_{i_1}, \dots, \epsilon_{i_K})' \qquad \text{ und } \qquad \epsilon = (\epsilon_{i_1}', \dots, \epsilon_{i_n}')'. \end{split}$$

Mit  $\beta = (\beta_1, \dots, \beta_K)'$  und der (K×K)-Matrix X, die in dem hier betrachteten Fall gleich der (K×K)-Einheitsmatrix I ist, lassen sich die Erwartungswertvektoren schreiben als

$$\mu_i := E(y_i \mid X\beta) = \Phi(-X\beta) = \Phi(-\beta)$$
 für alle  $i = 1,...,n$ .

112 Martin Spiess

Am Beginn des Kapitels wurde erwähnt, dass die Fähigkeiten  $\pi_{ik}$  als Funktion von Ladungen und latenten Variablen aufgefasst werden können. Bezeichnet I die geeignet dimensionierte Einheitsmatrix, die (K×J)-Matrix A die Matrix der Ladungen  $a_{kj}$  der k=1,...,K Items auf den j=1,...,J latenten Dimensionen ( $J \le K$ ) und  $\theta_i = (\theta_{i1},...,\theta_{iJ})'$  den J-dimensionalen Vektor der Fähigkeiten der i-ten Person auf den Dimensionen, mit  $\theta_i$  unabhängig und identisch verteilt nach N(0,1), dann ist

$$\pi_i = A\theta_i$$
 und  $(\pi_i + \varepsilon_i) \sim N(0, AA'+I - Diag(\sigma_1^2, ..., \sigma_K^2))$ ,

mit  $\sigma_1^2,\ldots,\sigma_K^2$  den Diagonalelementen der Matrix AA' und Diag $(\sigma_1^2,\ldots,\sigma_K^2)$  einer Diagonalmatrix mit Diagonalelementen  $\sigma_1^2,\ldots,\sigma_K^2$ . Damit ist AA'+I - Diag $(\sigma_1^2,\ldots,\sigma_K^2)$  – unter Restriktionen – eine positiv definite Korrelationsmatrix. Im nächsten Kapitel soll es allerdings lediglich um die Schätzung der Itemparameter gehen.

## 3 GEE-Schätzung der Itemparameter

Den GEE-Schätzer des obigen Modells erhält man im Allgemeinen durch iteratives Lösen der verallgemeinerten Schätzgleichungen

$$0 = \sum_{i=1}^{n} \frac{\partial \mu_i}{\partial \beta} W_i^{-1} (y_i - \mu_i)$$
(4)

nach dem interessierenden Parameter des systematischen Teils des Modells, im Falle des IRT-Modells  $\beta.$  Der (K×1)-Vektor  $\mu_i$  ist der korrekt spezifizierte Erwartungswert E(y\_i |  $\beta$ ), mit  $\beta$  den Itemschwierigkeiten. Zu beachten ist, dass hier der unbedingte und nicht der personenspezifische Erwartungswert modelliert wird. Weiterhin gelten die Schätzgleichungen sowie die folgenden Argumente für alle eindimensionalen Arten von Antwortvariablen, etwa stetige oder Zählvariablen. Schließlich ist

$$\frac{\partial \mu_i}{\partial \beta} = X'D'_i \qquad \text{mit} \qquad D_i = \text{Diag}(\phi(-\beta_1), \dots, \phi(-\beta_K))$$
 (5)

und  $\phi(\cdot)$  der Dichtefunktion der Standardnormalverteilung.  $W_i$  ist eine Gewichtungsmatrix

$$W_i = V_i^{1/2}RV_i^{1/2}$$
 mit  $V_i = Diag(var(y_{i1}), ..., var(y_{iK})),$ 

wobei im binären Modell  $var(y_{ik}) = \Phi(-\beta_k)(1-\Phi(-\beta_k))$  ist. R ist eine "Arbeits"-Korrelationsmatrix. Ist R die "wahre" Korrelationsmatrix für alle Personen, dann ist  $W_i$  gleich der Kovarianzmatrix  $Cov(y_i)$ . Die Lösung dieser Schätzgleichungen  $\hat{\beta}$  wird als GEE-Schätzer bezeichnet. Liang & Zeger (1986) zeigen, dass  $\hat{\beta}$  unter Annah-

men konsistent und asymptotisch normal verteilt ist, wenn der Erwartungswert korrekt spezifiziert ist und zwar selbst dann, wenn R fehlspezifiziert ist.

Liegen keine fehlenden Werte vor, d. h. beantworten alle Personen alle Items, dann ist

$$\mu_i = \mu$$
,  $V_i = V$   $W_i = W$  und  $D_i = D$  für alle i.

Die Schätzgleichungen werden zu

$$0 = (1' \otimes X')(I \otimes D')(I \otimes W)^{-1}(\gamma - 1 \otimes \mu), \tag{6}$$

wobei 1 den (n $\times$ 1)-Einservektor und  $\otimes$  das Kroneckerprodukt bezeichnet, und lassen sich vereinfachen zu

$$0 = X'D'W^{-1}\overline{y} - X'D'W^{-1}\mu, \quad \text{mit} \quad \overline{y} = n^{-1}\sum_{i} y_{i}.$$
 (7)

Da generell für jedes Item eine eigene Schwierigkeit geschätzt wird, ist X regulär und invertierbar. In dem hier besprochenen Modell ist darüber hinaus X = I. Die Schätzgleichungen vereinfachen sich weiter zu

$$0 = \overline{y} - \mu \quad \text{und damit} \quad \mu = \overline{y}, \tag{8}$$

denn auch D ist invertierbar. Da als Erwartungswertfunktion die Verteilungsfunktion der Standardnormalverteilung verwendet wurde, lässt sich der GEE-Schätzer ohne iteratives Verfahren direkt in einem Schritt berechnen:

$$\hat{\beta} = \Phi^{-1}(\overline{y}), \tag{9}$$

mit  $\Phi^{-1}(\cdot)$  der Inversen der Verteilungsfunktion der Standardnormalverteilung. Die Existenz von  $\hat{\beta}$  setzt also voraus, dass jedes Item von wenigstens einer bzw. von höchstens n-1 Personen korrekt beantwortet wird.

Nach Liang & Zeger (1986) ist  $\hat{\beta}$  unter schwachen Bedingungen konsistent und asymptotisch normal verteilt. Zur asymptotischen Effizienz ist bislang noch nichts ausgesagt. Der Varianzschätzer von  $n^{1/2}(\hat{\beta}-\beta_0)$  ist, mit Cov(y<sub>i</sub>) der wahren Kovarianzmatrix von y<sub>i</sub> und X = I,

$$Cov(n^{1/2}(\hat{\beta}-\beta_0)) =$$

$$\lim_{n\to\infty} n[(D'W^{-1}D)^{-1}(D'W^{-1}\{\frac{1}{n}\sum_{i}Cov(y_{i})\}W^{-1}D)(D'W^{-1}D)^{-T}]_{\beta_{0}}. \tag{10}$$

Unter Berücksichtigung der Symmetrie von D vereinfacht sich auch dieser Ausdruck, so dass

114 Martin Spiess

$$Cov(n^{1/2}(\hat{\beta} - \beta_0)) = \lim_{n \to \infty} n[D^{-1}\{\frac{1}{n}\sum_{i}Cov(y_i)\}D^{-1}]_{\beta_0}.$$
 (11)

McCullagh (1983) folgend besitzt  $\hat{\beta}$  minimale asymptotische Varianz innerhalb der Klasse von Schätzern mit linearer Einflussfunktion. Unter anderen gehören dazu auch Maximum-Likelihood-Schätzer verschiedener Modelle. Damit ist der GEE-Schätzer asymptotisch recht effizient. Zu beachten ist, dass dies unabhängig von der angenommenen Faktorenstruktur bzw. Dimensionalität gilt, solange lediglich das Erwartungswertmodell korrekt spezifiziert ist (siehe auch Spiess & Hamerle 1996).

Einen Schätzer für die Kovarianzmatrix von  $\hat{\beta}$  erhält man, indem alle Unbekannten durch deren Schätzer ersetzt werden,

$$\hat{V}(\hat{\beta}) = [D^{-1} \{ \frac{1}{n} \sum_{i} (y_{i} - \mu)(y_{i} - \mu)' \} D^{-1}]_{\hat{\beta}}.$$
 (12)

## 4 Erweiterungen

Die in Kapitel 3 beschriebene Vorgehensweise lässt sich ohne Einschränkungen auf IRT-Modelle mit anderen eindimensionalen Antworten erweitern, sofern die Erwartungswertfunktion und deren Inverse der Antwortvariablen analytisch verfügbar ist. Dies gilt für stetige Antworten, aber auch für Zählvariablen, wie etwa das Poissonmodell von Rasch (1960, 1961).

Im Falle ordinaler oder nominaler Antwortvariablen scheint eine so weit gehende Vereinfachung, wie im letzten Kapitel beschrieben, im Allgemeinen nicht möglich zu sein. Dennoch kann natürlich auch in diesen Fällen ein GEE-Schätzer berechnet werden. Zwar ist dann mit einem Effizienzverlust zu rechnen, dafür ist die Berechnung der Schätzwerte sehr einfach.

Häufig besteht über die Itemparameter hinaus auch Interesse an den Ladungen der Items. Um diese im Rahmen des GEE-Ansatzes schätzen zu können, bietet sich die in Spiess (2006) vorgeschlagene Erweiterung an. Dort werden die Schätzgleichungen für die Itemparameter des Erwartungswertmodells durch Schätzgleichungen für die Kovarianzstrukturparamater ergänzt. Benötigt werden dann zusätzlich zu den univariaten marginalen Verteilungsannahmen Modelle für alle möglichen bivariaten marginalen Verteilungen. Ein solcher Schätzansatz wäre für die betrachteten Item-Response-Modelle noch auszuarbeiten. Allerdings scheint eine Vereinfachung, wie im Falle der Itemparameterschätzung bei eindimensionalen Antwortvariablen, auch hier nicht möglich, so dass auf ein iteratives Verfahren zur Berechnung der Schätzer zurückgegriffen werden müsste.

Die Eigenschaften der Schätzer können im GEE-Ansatz in Anspruch genommen werden, wenn die korrekt spezifizierte marginale univariate Verteilung der Ant-

wortvariablen ein verallgemeinertes lineares Modell (generalized linear model) ist. Eine Verallgemeinerung ist mittels der verallgemeinerten Momentenmethode (generalized methods of moments; GMM) möglich (siehe Pannenberg & Spiess 2009). Schließlich sei noch erwähnt, dass das Erwartungswertmodell nicht auf die hier verwendete Normalverteilung beschränkt ist. Stattdessen kann die logistische Verteilung (Logitmodell) verwendet werden, wobei implizit angenommen wird, dass die Summe  $\nu_{ik}=\pi_{ik}+\epsilon_{ik}$  logistisch verteilt ist. Dies stimmt aber nicht mit der häufigen Annahme überein, dass  $\epsilon_{ik}$  logistisch verteilt ist. Bei der Anwendung des GEE-Ansatzes wird in diesem Fall entweder die Verteilung der Summe einer logistisch und einer davon unabhängig normal verteilten Zufallsvariable oder die Verteilung zweier unabhängig logistisch verteilter Zufallsvariablen benötigt. Sollen über die Itemparameter hinaus auch Ladungen geschätzt werden, werden zusätzlich die bivariaten marginalen Verteilungen benötigt.

### 5 Eine Simulationsstudie

In diesem Kapitel werden die Ergebnisse einer Simulationsstudie präsentiert, die Aufschluss über die Effizienz der vorgeschlagenen Schätzmethode geben soll. Mit Effizienz ist auch der rechnerische Aufwand und die Einfachheit der Methode gemeint. Gleichzeitig sollen die Simulationsergebnisse einen, wenn auch nur auf die simulierten Situationen beschränkten, Eindruck bezüglich der Eigenschaften der GEE-Schätzer liefern.

Das wahre datengenerierende Modell folgt dem in Kapitel 2 beschriebenen IRT-Modell. Der Einfachheit halber werden lediglich fünf Items betrachtet, die gelöst werden oder auch nicht (binäre Antwortvariable). Es wird weiterhin eine eindimensionale latente, individuenspezifische Variable erzeugt, so dass  $\pi_{ik}=\pi_i$  für alle k und  $\sigma_k^2=\sigma^2$ , mit  $\sigma^2=0.2$ ,  $\sigma^2=0.5$  und  $\sigma^2=0.8$ . Damit ist  $\text{var}(\epsilon_{ik})=0.8$ ,  $\text{var}(\epsilon_{ik})=0.5$  bzw.  $\text{var}(\epsilon_{ik})=0.2$  und die Korrelation der  $\nu_{ik}$  innerhalb der Personen 0.2, 0.5 bzw. 0.8.

Der in Kapitel 3 vorgeschlagene GEE-Schätzer wird verglichen mit dem MML-Schätzer eines generalisierten linearen gemischten Modells mit einer zufälligen latenten Variablen (generalized linear mixed model). Das zu berechnende Integral wird mit Hilfe der adaptiven Gauss-Hermite-Quadratur mit einer unterschiedlichen Anzahl an Stützstellen approximiert, wobei gelten sollte: Je mehr Stützstellen, desto besser die Approximation, aber desto länger auch die Rechenzeit. Verwendet wurde die Funktion glmer aus dem Ime4-Paket von R.

Um die Schätzer basierend auf den beiden Schätzverfahren vergleichen zu können, werden die GEE-Schätzer transformiert, denn bei der ML-Schätzung wird  $\beta_1^*,\ldots,\beta_k^*$  und bei der GEE-Schätzung  $\beta_1,\ldots,\beta_k$  geschätzt. Die wahren Itemparameter sind 0.5, 0.8, 1, 1.2 und 1.5. Die Simulationen wurden mit n = 100 und n = 500 durchgeführt. Als Anzahl an Stützstellen für die adaptive Gauss-Hermite-Quadratur wurden 5, 10 und 20 Punkte gewählt. Dies ergibt 18 Simulationsläufe, bei denen der

116 Martin Spiess

ML-Schätzer und sechs Simulationsläufe, bei denen der GEE-Schätzer berechnet wurde, wobei jeweils 500 Datensätze erzeugt wurden.

Um die Resultate vergleichen zu können, wurden folgende Kennwerte berechnet:

- der Mittelwert der Schätzwerte über die Simulationen (m),
- die (positive Quadrat-)Wurzel des Mittelwerts der geschätzten Varianzen über die Simulationen ( â ),
- die (positive Quadrat-)Wurzel der Varianz der Schätzwerte über die Simulationen (s),
- die mittlere Breite der Konfidenzintervalle mit einem  $\alpha = 0.05$  ( $\delta_{KI}$ ),
- der Anteil an Nicht-Ablehnungen der Nullhypothesen, bei der auf die wahren Parameter getestet wird ( $H_0$ ,  $\alpha=0.05$ ) und
- die CPU-Zeit für die jeweils 500 Simulationen in Stunden: Minuten: Sekunden (CPU).

	$n = 100, \sigma^2 = 0.2$									
				n = 100	1					
		Itemschwi	erigkeit: 0	.5	Itemschwierigkeit: 1.5					
	GEE	MML <sub>5</sub>	MML <sub>10</sub>	MML <sub>20</sub>	GEE	MML <sub>5</sub>	MML <sub>10</sub>	MML <sub>20</sub>		
m	0.508	0.518	0.516	0.510	1.530	1.529	1.533	1.536		
σ̂	0.021	0.020	0.020	0.020	0.042	0.041	0.042	0.042		
S	0.021	0.024	0.022	0.025	0.046	0.044	0.056	0.047		
$\delta_{\text{KI}}$	0.572	0.558	0.557	0.556	0.795	0.793	0.797	0.797		
H <sub>0</sub>	0.964	0.930	0.958	0.938	0.946	0.944	0.934	0.942		
CPU	0:00:01	0:10:45	0:14:44	0:21:52	0:00:01	0:10:45	0:14:44	0:21:52		
nsim	500	500	500	499	500	500	500	499		
				n = 500	$\sigma^2 = 0.2$					
		Itemschwi	erigkeit: 0	.5	Itemschwierigkeit: 1.5					
	GEE	MML <sub>5</sub>	MML <sub>10</sub>	MML <sub>20</sub>	GEE	MML <sub>5</sub>	MML <sub>10</sub>	MML <sub>20</sub>		
m	0.499	0.504	0.495	0.501	1.511	1.508	1.501	1.506		
σ̂	0.004	0.004	0.004	0.004	0.008	0.008	0.008	0.008		
S	0.005	0.004	0.005	0.004	0.008	0.008	0.008	0.009		
$\delta_{\text{KI}}$	0.255	0.246	0.247	0.246	0.348	0.346	0.345	0.346		
H <sub>0</sub>	0.942	0.938	0.930	0.926	0.960	0.946	0.960	0.946		
CPU	0:00:02	0:31:45	0:44:54	1:09:19	0:00:02	0:31:45	0:44:54	1:09:19		
nsim	500	500	500	500	500	500	500	500		

Tabelle 1: Simulationsergebnisse für  $\sigma^2=0.2$  mit n=100 bzw. n=500 Mittelwert (m), geschätzte Standardabweichung ( $\hat{\sigma}$ ), Standardabweichung (s), mittlere Breite der Konfidenzintervalle ( $\delta_{KI}$ ), Anteil an Nicht-Ablehnungen (H $_0$ ), CPU-Zeit in Stunden:Minuten:Sekunden (CPU), Anzahl erfolgreicher Simulationen (nsim). GEE-(GEE-) und MML-(ML-)Schätzer mit 5 (MML $_5$ ), 10 (MML $_{10}$ ) und 20 (MML $_{20}$ ) Stützstellen, 500 Simulationen.

	$n = 100, \sigma^2 = 0.5$									
		Itemschwi	ierigkeit: 0	.5	Itemschwierigkeit: 1.5					
	GEE	MML <sub>5</sub>	MML <sub>10</sub>	MML <sub>20</sub>	GEE	MML <sub>5</sub>	MML <sub>10</sub>	MML <sub>20</sub>		
m	0.513	0.500	0.516	0.521	1.536	1.522	1.553	1.552		
σ̂	0.033	0.029	0.030	0.030	0.050	0.046	0.047	0.057		
S	0.033	0.035	0.035	0.035	0.051	0.053	0.057	0.058		
$\delta_{\text{KI}}$	0.714	0.671	0.678	0.672	0.873	0.838	0.850	0.847		
H <sub>0</sub>	0.970	0.918	0.943	0.927	0.960	0.940	0.935	0.925		
CPU	0:00:01	0:12:11	0:16:43	0:24:00	0:00:01	0:12:11	0:16:43	0:24:00		
nsim	500	499	499	492	500	499	499	492		
				n = 500	$\sigma^2 = 0.5$					
		Itemschwi	ierigkeit: 0	.5	Itemschwierigkeit: 1.5					
	GEE	MML <sub>5</sub>	MML <sub>10</sub>	MML <sub>20</sub>	GEE	MML <sub>5</sub>	MML <sub>10</sub>	MML <sub>20</sub>		
m	0.507	0.507	0.504	0.501	1.506	1.518	1.516	1.513		
σ̂	0.007	0.005	0.006	0.006	0.010	0.008	0.008	0.009		
S	0.006	0.007	0.007	0.007	0.010	0.011	0.010	0.011		
$\delta_{\text{KI}}$	0.318	0.298	0.298	0.297	0.385	0.371	0.370	0.370		
H <sub>0</sub>	0.962	0.926	0.934	0.946	0.938	0.926	0.928	0.940		
CPU	0:00:02	0:36:07	0:51:48	1:17:53	0:00:02	0:36:07	0:51:48	1:17:53		
nsim	500	500	500	500	500	500	500	500		

Tabelle 2: Simulationsergebnisse für  $\sigma^2=0.5$  mit n=100 bzw. n=500 Mittelwert (m), geschätzte Standardabweichung ( $\hat{\sigma}$ ), Standardabweichung (s), mittlere Breite der Konfidenzintervalle ( $\delta_{KI}$ ), Anteil an Nicht-Ablehnungen (H $_0$ ), CPU-Zeit in Stunden:Minuten:Sekunden (CPU), Anzahl erfolgreicher Simulationen (nsim). GEE-(GEE-) und MML-(ML-)Schätzer mit 5 (MML $_5$ ), 10 (MML $_{10}$ ) und 20 (MML $_{20}$ ) Stützstellen, 500 Simulationen.

Da die Konvergenzkriterien der ML-Schätzer nicht für alle Datensätze erfüllt waren, ist in den Tabellen 1-3 in einer zusätzlichen Zeile die jeweilige Anzahl an erfolgreichen Simulationen angegeben (nsim). Bei 500 erfolgreichen Simulationen sollte der Anteil an Nicht-Ablehnungen der  $H_0$  bei einem  $\alpha = 0.05$  etwa zwischen 0.93 und 0.97 liegen, wenn alle Annahmen erfüllt sind.

Aus Platzgründen werden in den Tabellen 1-3 lediglich die Ergebnisse für die wahren Itemschwierigkeiten 0.5 und 1.5 präsentiert.

Ein Unterschied zwischen GEE- und MML-Schätzung ist bereits auf einer rein technischen Ebene sehr auffällig: Der GEE-Schätzer ist, wie zu erwarten war, sehr schnell zu berechnen. Während die CPU-Zeiten bei 500 Simulationen für die Berechnung des GEE-Schätzers zwischen einer bzw. zwei Sekunden schwanken, wer-

118 Martin Spiess

	$n = 100, \sigma^2 = 0.8$								
		Itemschwi	erigkeit: 0	.5	Itemschwierigkeit: 1.5				
	GEE	MML <sub>5</sub>	MML <sub>10</sub>	MML <sub>20</sub>	GEE	MML <sub>5</sub>	MML <sub>10</sub>	MML <sub>20</sub>	
m	0.503	0.508	0.541	0.540	1.526	1.548	1.558	1.567	
σ̂	0.081	0.064	0.065	0.065	0.094	0.076	0.078	0.078	
S	0.085	0.099	0.114	0.118	0.098	0.119	0.142	0.149	
$\delta_{\text{KI}}$	1.112	0.984	0.996	0.997	1.202	1.078	1.087	1.090	
H <sub>0</sub>	0.934	0.896	0.877	0.848	0.950	0.873	0.856	0.848	
CPU	0:00:01	0:17:34	0:22:51	0:32:17	0:00:01	0:17:34	0:22:51	0:32:17	
nsim	500	482	471	500	500	482	471	500	
				n = 500	$\sigma^2 = 0.8$				
		Itemschwi	erigkeit: 0	.5	Itemschwierigkeit: 1.5				
	GEE	MML <sub>5</sub>	MML <sub>10</sub>	MML <sub>20</sub>	GEE	MML <sub>5</sub>	MML <sub>10</sub>	MML <sub>20</sub>	
m	0.499	0.486	0.497	0.489	1.508	1.483	1.493	1.489	
σ̂	0.016	0.012	0.012	0.012	0.019	0.014	0.015	0.015	
S	0.017	0.019	0.020	0.019	0.017	0.021	0.021	0.022	
$\delta_{\text{KI}}$	0.496	0.432	0.433	0.433	0.535	0.471	0.472	0.472	
H <sub>0</sub>	0.950	0.886	0.870	0.884	0.952	0.900	0.896	0.898	
CPU	0:00:02	0:52:33	1:13:18	1:51:10	0:00:02	0:52:33	1:13:18	1:51:10	
nsim	500	498	500	500	500	498	500	500	

Tabelle 3: Simulationsergebnisse für  $\sigma^2=0.8$  mit n=100 bzw. n=500 Mittelwert (m), geschätzte Standardabweichung ( $\hat{\sigma}$ ), Standardabweichung (s), mittlere Breite der Konfidenzintervalle ( $\delta_{KI}$ ), Anteil an Nicht-Ablehnungen ( $H_0$ ), CPU-Zeit in Stunden:Minuten:Sekunden (CPU), Anzahl erfolgreicher Simulationen (nsim). GEE-(GEE-) und MML-(ML-)Schätzer mit 5 (MML $_5$ ), 10 (MML $_{10}$ ) und 20 (MML $_{20}$ ) Stützstellen, 500 Simulationen.

den für die Berechnung der MML-Schätzer zwischen ca. 10 Minuten und etwa 1 Stunde und 51 Minuten benötigt. Die Rechenzeit hängt von der Anzahl an Personen und der Anzahl an gewählten Stützstellen ab. Zu berücksichtigen ist, dass für eine Inanspruchnahme der für ML-Schätzer gezeigten günstigen Eigenschaften sowohl n als auch die Anzahl an Stützstellen gegen unendlich gehen muss. Das bedeutet, dass bei einer zu geringen Anzahl an Stützstellen auch nicht zu erwarten ist, dass der MML-Schätzer oder dessen Varianzschätzer selbst in großen Stichproben approximativ unverzerrt ist.

Betrachtet man die eher "inhaltlichen" Ergebnisse, so fällt zunächst auf, dass die mittleren Schätzwerte, sowohl des GEE- als auch des MML-Schätzers, im Allgemeinen nahe an den wahren Werten liegen. Während der GEE-Schätzer unproblematisch zu sein scheint – die mittleren Varianzschätzer und die Varianzen der Schätz-

werte über die Simulationen sind sehr ähnlich – unterschätzt der MML-Schätzer die tatsächliche Varianz im Bereich der niedrigen und mittleren Korrelation. Dies drückt sich in kürzeren mittleren Konfidenzintervallen, aber auch in etwas zu niedrigen Ablehn- oder Überdeckungsraten aus (vgl. die Tab. 1 und 2).

Sind die Schätzergebnisse im Bereich der niederen bis mittleren Korrelationen für die ML-Schätzer teilweise bereits grenzwertig, so sind die Anzahl an nicht erfolgreichen Simulationen sowie die Ablehnraten der Nullhypothese bei hoher Korrelation (Bedingung  $\text{var}(\epsilon_{ik}) = 0.8$ ) nicht mehr akzeptabel. Dabei hilft ganz offensichtlich auch keine Erhöhung der Anzahl an Stützstellen auf 20. Möglicherweise könnten die Schätzergebnisse für den ML-Schätzer bei einer noch höheren Anzahl an Stützstellen verbessert werden, bei einer CPU-Zeit von knapp zwei Stunden für die Simulation mit n = 500 Einheiten und 20 Stützstellen lässt sich aber absehen, dass entsprechende Simulationen enorm zeitaufwendig wären, zumal ja die für eine akzeptable Schätzung notwendige Anzahl an Stützstellen unbekannt ist und daher die Anzahl sukzessive zu erhöhen wäre.

### 6 Diskussion

Der vorgeschlagene GEE-Schätzer steht analytisch, in einem Schritt zur Verfügung. Er existiert im Modell mit binären Antwortvariablen dann, wenn jedes Item von wenigstens einer Person bzw. höchstens n-1 Personen gelöst wird. Der GEE-Schätzer ist konsistent, asymptotisch normal verteilt und besitzt asymptotisch eine vergleichsweise kleine Varianz. Zumindest was die Schätzung der Schwierigkeitsparameter von Items angeht, ist der GEE-Schätzer damit den gängigen ML-Schätzern in vielerlei Hinsicht überlegen. Die in Kapitel 5 dargestellten Ergebnisse illustrieren zumindest für die hier simulierten Situationen, dass der vorgeschlagene GEE-Schätzer im Vergleich zu dem verwendeten ML-Schätzer nicht nur sehr stabil und schnell zu berechnen ist, sondern darüber hinaus auch in endlichen Stichproben effizient zu sein scheint.

Mit diesen Ergebnissen ist das vorgeschlagene Verfahren bereits eine attraktive Alternative zu den gängigen Vorgehensweisen. Zu beachten ist, dass dies nicht nur für die Annahme eines Modells mit einer wie im Simulationsteil betrachteten einfachen Kovarianzstruktur, sondern erheblich allgemeiner, für beliebige positiv definite Kovarianzstrukturen gilt. Von Interesse wäre eine Erweiterung auch auf ordinale bzw. nominale Modelle. Eine andere interessante Erweiterung besteht in der Ableitung der Schätzgleichungen für die Itemladungen und in der Frage, inwieweit hier eine Vereinfachung gegenüber den gängigen Verfahren erreicht werden kann.

## Das Beck Depressions-Inventar (BDI-II).

# **Psychometrische Analysen mit probabilistischen Testmodellen**

Ferdinand Keller

## 1 Entwicklung des BDI und bisherige psychometrische Befunde

Das Beck Depressions-Inventar (Beck et al. 1961) ist ein Selbstbeurteilungsinstrument zur Erfassung der Schwere depressiver Symptomatik. Es wird seit nunmehr fast 50 Jahren im englischen Sprachraum eingesetzt und hat Übersetzungen in alle Kultursprachen der Welt erfahren. Seit fast 40 Jahren liegt eine deutsche Übersetzung des BDI vor, die nach mehreren Überarbeitungen und veränderten Vorschlägen schließlich in eine verbindliche Form gebracht wurde (Hautzinger et al. 1994). Nach mehreren Jahrzehnten Erfahrung mit dem ursprünglichen BDI und einem ergänzten BDI-IA (Beck & Steer 1987) legten Beck et al. (1996) eine Revision vor, das BDI-II. Hauptanliegen war eine Verbesserung der inhaltlichen Validität des Verfahrens durch die Aufnahme neuer diagnoserelevanter und den Ausschluss diagnoseirrelevanter Items.

Die ursprüngliche Version des BDI beruhte auf typischen Aussagen über Symptome, die häufig von psychiatrischen Patienten mit Depression und selten von solchen ohne gemacht wurden. Dies führte zu einer Zusammenstellung von 21 Items, die depressionscharakteristische Symptome und Einstellungen erfassen sollten. Die Items waren durch eine vierstufige Antwortskala (0-3) mit inhaltlich ausformulierten Aussagen repräsentiert, die eine aufsteigende Schwere der Symptomatik abbilden. In der revidierten Form des BDI (BDI-II) wurden insgesamt vier Items (Gewichtsverlust, Veränderung des Körperbilds, intensive Beschäftigung mit körperlichen Symptomen, Arbeitsschwierigkeiten) eliminiert und durch vier neue (Unruhe, Gefühl der Wertlosigkeit, Konzentrationsschwierigkeiten und Energieverlust) ersetzt. Zwei Items wurden so abgeändert, dass sie nun sowohl die Zunahme als auch die Abnahme von Appetit und Schlaf erfassen. Darüber hinaus wurden zahlreiche Aussagen zu anderen Symptombereichen neu formuliert. Schließlich wurde in Anlehnung an die aktuellen psychiatrischen Diagnosesysteme (DSM-IV, APA 1994; ICD-10, WHO 1992) das Zeitkriterium zur depressiven Symptomatik auf die zurückliegenden zwei Wochen erweitert. Die Revision des BDI ging somit mit zahlreichen inhaltlichen und formalen Veränderungen einher, was notwendigerweise auch die Frage nach den psychometrischen Gütekriterien neu stellte.

Im Folgenden wird daher nur kurz und überblicksartig auf die Ergebnisse zur Skalen- bzw. Subskalenanalyse der ersten Version des BDI eingegangen, bevor über die Darstellung der Ergebnisse bisheriger psychometrischer Analysen zum BDI-II die Fragestellungen dieses Beitrags abgeleitet werden, der insbesondere die vergleichende Analyse verschiedener Stichproben mit probabilistischen Testmodellen behandelt.

### 1.1 Befunde zum BDI

Bereits mit der ursprünglichen Version des BDI wurden zahlreiche Studien zur faktoriellen Validität durchgeführt. Beck et al. (1988) resümieren, dass je nach Extraktions- und Abbruchkriterium drei bis sieben Faktoren identifiziert wurden. Für dessen deutsche Fassung ließen sich in ähnlicher Weise methodenspezifisch unterschiedlich viele Faktoren identifizieren, wobei jedoch im Wesentlichen nur ein genereller Faktor replizierbar war (Richter et al. 1994, Hautzinger et al. 1994, Richter et al. 1998).

Psychometrische Analysen mit probabilistischen Testmodellen sind selten geblieben. Bouman & Kok (1987) verwendeten das dichotome Rasch-Modell und ihre Itemanalysen des BDI zeigen, dass nicht von einer Unidimensionalität der Skala ausgegangen werden kann. Stattdessen identifizierten sie drei verschiedene, in sich homogene Subskalen, die sie mit Guilt & Failure, Mood & Inhibition und Somatic bezeichneten. Eine Replikation dieser Subskalen in einer deutschen Stichprobe erbrachte jedoch keine Homogenität im Sinne des Rasch-Modells, wenn auch die qualitativen Abweichungen von Antwortmustern gering waren und deshalb auch für die deutsche Version des BDI von zufrieden stellenden psychometrischen Eigenschaften ausgegangen werden konnte (Keller & Kempf 1993). Weitere Analysen mit neuen, inhaltlich zusammengestellten Subskalen und unter Verwendung der vier Kategorien je Item erbrachten ähnliche Ergebnisse: Die strengen Annahmen der Rasch-Homogenität waren zwar meist nicht erfüllt, ihre Verletzungen aber eher geringfügig und meist auf einzelne Items zurückzuführen, deren Kategorienformulierungen teilweise missverständlich und deren Schwellenparameter folglich ungeordnet waren (Keller & Kempf 1997). Für eine kognitive Subskala, bestehend aus sechs Items, ergab sich allerdings Rasch-Homogenität, wenn der BIC als Kriterium herangezogen wurde (Keller 1998, 2000).

### 1.2 Befunde zum BDI-II

Für die revidierte Fassung (BDI-II) stellt sich eine ähnliche Situation dar. Auch dieses erweist sich als reliables und valides Testinstrument (Beck et al. 1996), was ebenso für die deutsche Version bestätigt werden kann (Kühner et al. 2007). Bezüglich der faktoriellen Struktur ergibt sich allerdings im Unterschied zu den älteren Analysen, dass die explorativen und konfirmatorischen Faktorenanalysen eine Aufsplittung des BDI-II in einen kognitiven und einen somatisch-affektiven Faktor nahe legen (Steer et al. 1999). Für das deutschsprachige BDI-II konnte diese

122 Ferdinand Keller

Struktur ebenfalls bestätigt werden (Keller et al. 2008). In mehreren Studien ließ sich ein Modell mit einem generellen Depressionsfaktor sowie einem somatischen und einem kognitiven Faktor als am besten passend identifizieren, wobei trotzdem der BDI-II Gesamtwert benutzt werden kann (Quilty et al. 2010).

Wie beim BDI stehen auch beim BDI-II zahlreiche faktorenanalytische Publikationen nur einer sehr kleinen Anzahl von Publikationen gegenüber, in denen probabilistische Testmodelle verwendet wurden. Beck et al. (1996) benutzten das Programm TESTGRAF, das bereits früher angewendet wurde (Santor et al. 1994) und eine nonparametrische Itemanalyse gestattet. Im Manual der deutschen Version (Hautzinger et al. 2006) wird das ordinale Rasch-Modell verwendet.

Zusammenfassend ergab sich in beiden Analysen für die Gruppe der Depressiven, dass die Schwellenwerte bei den meisten Items die erwartete ordinale Abstufung aufweisen und damit eine anwachsende Beziehung zur gesamten (latenten) Depressivität gegeben ist. Einige Items folgen jedoch nicht der antizipierten Reihenfolge, insbesondere Item 6 (Bestrafungsgefühle). Kategorie 3 ist hier sogar die leichteste und wird schon bei relativ niedriger Depressivität angekreuzt. Die Kategorienformulierungen dieses Items werden auch bei Beck et al. (1996) als problematisch identifiziert und die Items 9 (Suizidgedanken), 11 (Unruhe) und 21 (Verlust an sexuellem Interesse) werden als nicht monoton bezeichnet. Alle drei erscheinen in der deutschen Stichprobe geordnet. Insbesondere für Suizidalität ergibt sich eine ideale Abstufung, die bei Beck et al. vermutlich wegen zu geringer Besetzungen der Kategorien 2 und 3 nicht analysiert werden konnte. Problematisch erscheint noch das Item 10 (Weinen), das kaum diskriminiert.

Trotz dieser Einschränkungen ergibt sich gemäß ordinalem Rasch-Modell eine hohe Reliabilität des BDI-II; sie liegt bei 0.92. In den anderen Stichproben im deutschen Manual (Gesunde und eine Internet-Stichprobe) fällt die Reliabilität etwas ab (0.80 bzw. 0.82).

Angesichts der zunehmenden Verbreitung des BDI-II über den klinischen Bereich hinaus erscheint es interessant zu analysieren, wie gut es bei Gesunden misst. Auch die psychometrischen Eigenschaften bei nicht primär depressiven Patienten sind noch unzureichend untersucht.

Eine Besonderheit des BDI-II liegt darin, dass die einzelnen Schwierigkeitskategorien für ein Symptom mit einer inhaltlichen Aussage formuliert sind und damit nicht die meist übliche Kategorienbezeichnung nach Intensität oder Häufigkeit übernommen wird, die für alle Items einer Skala identisch ist (z. B. "selten", "manchmal", "öfter", "meistens"). Daher kann nicht nur geprüft werden, ob die Kategorien im Verständnis der Probanden eine aufsteigende Reihenfolge in der Depressivität abbilden, sondern auch, wie sich verschiedene Modellannahmen über Kategorienabstände auswirken.

## 1.3 Fragestellungen dieses Beitrags

Vier teilweise zusammenhängende Fragenkomplexe werden bearbeitet:

- 1. Wie gut misst das BDI-II Depressivität in verschiedenen Stichproben? Da unter Gültigkeit des Rasch-Modells spezifische Objektivität gegeben sein müsste, sollten die Itemparameter in allen Stichproben gleich sein.
- 2. Ordinalität der Antwortkategorien: Durch die Ausformulierung der Antwortkategorien in den BDI-Items ist zu prüfen, ob sie die erwartete ordinale Abstufung aufweisen und damit eine anwachsende Beziehung zur gesamten (latenten) Depressivität gegeben ist. Weiterhin können zusätzliche Annahmen über die Schwellenabstände und folglich die Anzahl notwendiger Schwellenparameter getestet werden.
- 3. In Erweiterung der beiden vorigen Punkte: Gilt das Rasch-Modell für die gesamte Personenstichprobe oder gilt es mit jeweils unterschiedlichen Modellparametern nur in verschiedenen, unbekannten Teilstichproben? Diese Verallgemeinerung wird als Mixed-Rasch-Modell bezeichnet (Rost 2004).
- 4. Aus konfirmatorischen Faktorenanalysen (CFA) ist bekannt, dass für das BDI-II eine zweidimensionale Struktur anzunehmen ist, wobei die beiden Dimensionen hoch korreliert sind. Insbesondere eine kognitive Subskala zeichnet sich konsistent ab (Keller et al. 2008). Da das Rasch-Modell die Unidimensionalität der Skala voraussetzt bzw. prüft und damit ein theoretisches Problem entsteht, sollen die Analysen auch für die kognitive Subskala durchgeführt werden.

### 2 Methoden

# 2.1 Stichproben

Die Analysen entstanden in Auswertung von drei Stichproben<sup>1</sup>. Die ersten beiden stammen aus dem deutschen BDI-II Manual (Hautzinger et al. 2006) und sind dort ausführlicher beschrieben.

Stichprobe 1: Depressive Patienten (n = 266). Es handelt sich um unipolar depressive Patienten, die in verschiedenen Studienzentren (Mannheim, Tübingen, Bayreuth, Ulm, Lübeck, Kiel und Weissenau) in ambulanter oder stationärer Behandlung waren. Das Durchschnittsalter der Probanden betrug 48.8 Jahre (SD = 15.7, Range 18-85), der Frauenanteil lag bei 65.4%.

Stichprobe 2: Gesunde (n = 582). Diese Stichprobe umfasst 582 Personen aus der Allgemeinbevölkerung, die parallel zur o. g. Patientenstichprobe untersucht wurden. Das Durchschnittsalter betrug hier 36.2 Jahre (SD = 15.1, Range 15-81), der Frauenanteil lag bei 65.6%.

<sup>1</sup> Ich danke Christine Kühner und Martin Hautzinger für die freundliche Zustimmung zur weiteren Auswertung der Manualstichproben und ganz besonders Herrn Dr. Robert Mestel, Abteilungsleiter Forschung/Qualitätssicherung der HELIOS Klinik Bad Grönenbach, für die Überlassung der Daten von Stichprobe 3.

124 Ferdinand Keller

Stichprobe 3: Patienten einer psychosomatischen Klinik (HELIOS Klinik Bad Grönenbach). Diese Stichprobe besteht aus 898 Patienten, die zwischen Herbst 2009 und Frühjahr 2010 stationär behandelt wurden. Die meisten Patienten (63.4%) weisen Diagnosen aus dem depressiven Formenkreis auf (Kapitel F3 des ICD 10), weitere 13.6% aus Kapitel F4 (Angst, PTBS, Somatisierung) und 12.1% aus Kapitel F6 (Persönlichkeitsstörungen). Lässt man eine F3-Diagnose auch als Komorbidität zu, weisen insgesamt 88.3% eine depressive Störung auf. Das mittlere Alter lag bei 44.0 Jahren (SD = 10.5; range: 18-71 Jahre); der Frauenanteil bei 66.7%.

### 2.2 Testmodelle

Die Analysen zum Rasch-Modell und dem Mixed-Rasch-Modell wurden mit dem Programm WINMIRA 2001 v1.45 von v. Davier (2001) durchgeführt, das die Berechnung von vier verschiedenen Modellen zur Parametrisierung der Antwortkategorien ermöglicht. Im Partial-Credit-Modell (PCM) wird angenommen, dass die Schwellenparameter für jedes Item unterschiedlich und klassenabhängig verschieden sind (Rost 2004). Die anderen drei Modelle setzen Restriktionen und reduzieren damit die notwendige Zahl der Parameter. Bei Skalen, in denen für jedes Item dieselben Antwortkategorien verwendet werden, könnte man annehmen, dass alle Items dieselben Schwellenabstände haben. Modelle, die auf dieser Annahme basieren, heißen Ratingskalen-Modelle (RSM – Rost 2004). Der Vergleich von PCM zu RSM ist daher besonders interessant, da das BDI-II eben nicht auf gleichen Antwortkategorien pro Item basiert, sondern jedes symptomspezifisch ausformuliert ist. Eine Bevorzugung des PCM spräche dann für die Notwendigkeit der Berücksichtigung dieser itemspezifischen Formulierungen, während das RSM auf letztlich doch itemübergreifende gleiche Kategorienabstände hinweisen würde. Als weitere Modelle sollen noch das Äquidistanz-Modell (die Kategorien haben gleiche Schwellenabstände innerhalb jedes Items, diese dürfen jedoch für jedes Item unterschiedlich groß sein) und das Dispersions-Modell, das eine Kombination aus den beiden letzteren darstellt (vgl. ebd., Kap. 3.3.4). Zur sparsameren Modellierung kann in WINMIRA eine Verteilungsapproximation verwendet werden ("smoothed" scores), die ebenfalls für die jeweils am besten passenden Modelle berechnet wird; die Qualität der Anpassung (WINMIRA berechnet dazu das RMSEA, vgl. z. B. Reinecke 2005) wurde dabei jedoch nicht im Einzelnen bewertet. Die Auswahl des "besten" Modells erfolgt anhand des BIC, es werden der Vollständigkeit halber aber auch AIC und CAIC in den Tabellen berichtet. Zur Überprüfung der Annahme, dass die Itemparameter in verschiedenen Stichproben homogen sind, wurde der Andersen-Test (vgl. Rost 2004) benutzt.<sup>2</sup>

<sup>2</sup> Verwendet dazu wurde das Modul eRm (Mair & Hatzinger 2007); mein herzlicher Dank geht an Rainer Alexandrowicz für die Durchführung dieser Berechnungen.

### 2.3 BDI-II

### 2.3.1 Inhalte der Items

Die 21 Items erfragen in vier Kategorien die Beschwerden und Einschränkungen in den folgenden Symptombereichen:

- Traurigkeit
- Pessimismus
- 3. Versagensgefühle
- 4. Verlust an Freude
- 5. Schuldgefühle
- 6. Bestrafungsgefühle
- 7. Selbstablehnung
- Selbstkritik
- 9. Suizidgedanken
- 10. Weinen
- 11. Unruhe
- 12. Interessenverlust
- 13. Entschlusslosigkeit
- 14. Wertlosigkeit
- 15. Energieverlust
- 16. Schlaf
- 17. Reizbarkeit
- 18. Appetitveränderung
- 19. Konzentrationsschwierigkeiten
- 20. Ermüduna
- 21. Verlust an sexuellem Interesse

Zur Verdeutlichung des Antwortformats werden beispielhaft die Antwortkategorien von zwei BDI-Items wiedergegeben:

### Traurigkeit

- 0 Ich bin nicht traurig.
- 1 Ich bin oft traurig.
- 2 Ich bin ständig traurig.
- 3 Ich bin so traurig oder unglücklich, dass ich es nicht aushalte.

#### 15. Energieverlust

- 0 Ich habe so viel Energie wie immer.
- 1 Ich habe weniger Energie als sonst.
- 2 Ich habe so wenig Energie, dass ich kaum noch etwas schaffe.
- 3 Ich habe keine Energie mehr, um überhaupt noch etwas zu tun.

#### 2.3.2 Faktorielle Struktur des BDI-II und Subskalen

Die Item-Faktor-Zuordnungen aus explorativen Faktorenanalysen wie auch bisherigen Befunden aus der Literatur (vgl. Keller et al. 2008) zeigen, dass die BDI-Items 2-3, 5-9 und 14 gut einem kognitiven Faktor zuordenbar sind, während die restli-

126 Ferdinand Keller

chen Items einen zweiten, somatisch-affektiven Faktor bilden. Ein "Grenzgänger" ist Item 1 (Traurigkeit), das auf beiden Faktoren etwa gleich lädt und abhängig von der Stichprobe (Patienten, Gesunde mit breitem Altersrange, Studenten) variieren kann, in der theoretischen Zuordnung aber eher zum Faktor somatisch-affektiv gezählt wird. Beide Faktoren korrelieren mit ca. r=0.60 miteinander (manifeste Ebene), so dass gemeinhin von einem generellen depressiven Syndrom ausgegangen wird, das in zwei hoch korrelierte Faktoren zerlegt werden kann.

### Kurzskala

Als trennscharfe Kurzversion bzw. für Screening-Zwecke bei somatischen Patienten oder bei Alkoholikern wurden für die ursprüngliche BDI-Version die drei Kernitems 4 (Verlust an Freude), 12 (Interessensverlust) und 13 (Entschlussunfähigkeit) vorgeschlagen (Clark et al. 1993; Keller et al. 2002). Auch in der revidierten Fassung ergeben sich für diese 3-Item-Skala sehr gute psychometrische Eigenschaften. Bei den Depressiven beträgt Cronbachs  $\alpha$  0.84, bei den Gesunden 0.70. Die Korrelation zum Gesamtwert des BDI-II liegt bei 0.89 (Depressive) bzw. 0.78 (Gesunde). Im Folgenden soll diese Kurzskala hinsichtlich ihrer Vergleichbarkeit der Itemparameter in verschiedenen Stichproben untersucht werden.

## 3 Ergebnisse

## 3.1 Vergleich der Rasch-Modelle innerhalb der einzelnen Stichproben

# 3.1.1 Depressive Patienten

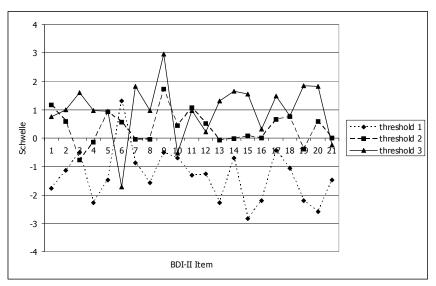


Abbildung 1: Schwellenwerte der BDI-II-Items bei depressiven Patienten (n = 266)

In der Depressionsstichprobe zeigt sich bei Betrachtung der BICs, dass die 1-Klassen-Lösung immer besser ist als eine 2-Klassen-Lösung (vgl. Tab. 1). Damit wird die Geltung des Rasch-Modells nahe gelegt. Beim Vergleich der unterschiedlichen Modellannahmen bezüglich der Kategorienabstände zeigt sich, dass das PCM den besten BIC aufweist. Obwohl es die meisten Parameter benötigt und im BIC dafür "bestraft" wird, scheinen alle diese unterschiedlichen Schwellen pro Item nötig zu sein, denn die restriktiveren Modelle, v. a. das RSM mit der Annahme gleicher Abstände pro Item, schneiden deutlich schlechter ab (vgl. Abb. 1). Unter der Annahme einer geglätteten Verteilung (smoothed) ergibt sich eine weitere Verbesserung im BIC des PCM; auch hier wird das 1-RM bevorzugt.

Rasch-Modell (RM)	Anzahl Parameter	BIC	CAIC	AIC	Reliabili- tät	Klassen- größe
Partial-Credit-Modelle						
1-RM	125	12130.2	12255.2	11682.2	.92	100
2-RM	249	12453.0	12702.0	11560.7	.92/.93	58 / 42%
1-RM smoothed	64	11850.4	11914.4	11621.0	.92	100
2-RM smoothed	129	11897.3	12026.3	11435.0	.91/.91	56 / 44%
Ratingskalen-Modelle						
1-RM	85	12572.0	12657.0	12267.4	.92	100
2-RM	169	12596.4	12765.4	11990.8	.92/.90	64 / 36%
Äquidistanz-Modelle						
1-RM	104	12192.3	12296.3	11819.7	.92	100
2-RM	207	12552.6	12759.6	11810.8	.91/.91	57 / 43%
Dispersions-Modelle						
1-RM	105	12139.4	12244.4	11763.2	.92	100
2-RM	209	12488.9	12697.9	11740.0	.92/.93	57 / 43%

Tabelle 1: Vergleich der Modellanpassung für verschiedene Rasch-Modelle in der Stichprobe Depressiver (n = 266)

# 3.1.2 Psychosomatik-Patienten

Für die Psychosomatikstichprobe ist im Unterschied zur Depressivenstichprobe jeweils eine 2-Klassen-Lösung vorzuziehen (vgl. Tab. 2). Damit wird Heterogenität nahe gelegt. Beim Vergleich der unterschiedlichen Modellannahmen bezüglich der Kategorienabstände zeigt sich wiederum, dass das PCM den besten BIC aufweist, d. h. auch in dieser Stichprobe scheinen alle diese unterschiedlichen Schwellen pro Item nötig zu sein, auch in den 2-Klassen-Lösungen. Unter der Annahme smoothed ergibt sich eine weitere Verbesserung im BIC des PCM; hier wird sogar das 3-RM bevorzugt.

128 Ferdinand Keller

Inhaltlich zeigt sich bei Betrachtung der 2-Klassen-Lösung eine Aufteilung in eine Gruppe mit eng beieinander liegenden Schwellenparametern und eine mit eher weit auseinander liegenden. Diese Muster können als response sets interpretiert werden und beschreiben eine Gruppe mit der Tendenz zu extremen Antworten (58%) und eine Gruppe mit Tendenz zur Mitte (42%).

Dasch Modell (DM)	Anzahl	BIC	CAIC	AIC	Reliabilität	Klassengröße
, ,	Parameter					
Partial-Credit-Modelle						
1-RM	125	41554.1	41679.1	40954.1	.91	100
2-RM	249	41349.7	41598.7	40154.5	.91/.91	58/42%
3-RM	373	41604.6	41977.6	39814.2	.90/.92/.88	41/31/28
1-RM smoothed	64	41193.3	41257.3	40886.1	.91	100
2-RM smoothed	129	40651.0	40780.0	40031.7	.91/.91	58/42%
3-RM smoothed	194	40545.6	40739.6	39614.4	.91/.85/.87	47/26/27%
4-RM smoothed	259	40578.2	40837.2	39334.9	88/87/86/87	33/24/22/21
Ratingskalen-Modelle						
1-RM	85	42616.0	42701.0	42208.0	.91	100
2-RM	169	42252.9	42421.9	41441.7	.92/.90	51/49%
Äquidistanz-Modelle						
1-RM	104	42024.6	42128.6	41525.4	.91	100
2-RM	207	41984.0	42191.0	40990.4	.90/.91	55/45%
Dispersions-Modelle						
1-RM	105	41863.6	41968.6	41359.6	.91	100
2-RM	209	41933.8	42142.8	40930.5	.91/.93	56/44%

Tabelle 2: Vergleich der Modellanpassung für verschiedene Rasch-Modelle in der Stichprobe Psychosomatik (n = 898)

#### **3.1.3** Gesunde

In der Stichprobe der Gesunden zeigt sich, dass die 1-Klassen-Lösung vorzuziehen ist (vgl. Tab. 3), womit die Geltung des Rasch-Modells nahe gelegt wird. Allerdings sind die Schwellenparameter zu Kategorie 3 teilweise nur sehr ungenau zu schätzen, da diese meist nur gering besetzt ist (bei 15 Items < 10 Nennungen, bei 11 Items sind es  $\leq$  5 Nennungen). Bei der Schätzung von Mixed-Rasch-Modellen wirkt sich dies noch stärker aus, weshalb auf die Berechnung der weiteren Modelle verzichtet wurde. Beim Vergleich der unterschiedlichen Modellannahmen bezüglich der Kategorienabstände zeichnet sich aber wie bei den anderen beiden Stichproben ab, dass das PCM den besten BIC aufweist.

Rasch-Modell (RM)	Anzahl Parameter	BIC	CAIC	AIC	Reliabili- tät	Klassen- größe
Partial-Credit-Modelle						
1-RM	125	15379.4	15504.4	14833.6	.80	100
2-RM	249	15786.7	16035.7	14699.4	.85/.73	63/37%
1-RM smoothed	64	15037.5	15101.5	14758.0	.80	100
2-RM smoothed	129	15174.3	15303.3	14611.0	.72/.88	62/38%
Ratingskalen-Modelle						
1-RM	85	16028.8	16113.8	15657.7	,79	100
2-RM	169	16060.8	16229.8	15322.9	.77/.83	66/34%

Tabelle 3: Vergleich der Modellanpassung für verschiedene Rasch-Modelle in der Stichprobe Gesunder (n = 582)

## 3.2 Vergleich der Itemparameter über die drei Stichproben

Die Itemparameter in allen drei Stichproben sind in Abb. 2 dargestellt, in der für jede Stichprobe das Rasch-Modell (1-Klassen-Lösung des PCM) unterstellt wurde. Weiterhin ist ein "Konfidenzband" abgebildet, das für die Depressionsstichprobe berechnet wurde, indem der Standardfehler pro Itemparameter mit 2 multipliziert wurde, so dass sich ein ungefähres 95%-Konfidenzintervall ergibt.

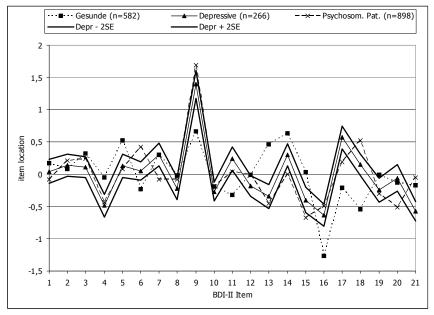


Abbildung 2: Vergleichende Darstellung der Itemparameter in den drei Stichproben

130 Ferdinand Keller

Beim Vergleich der Patientenstichproben (1 und 3) zeigt sich, dass bei vielen Items die Itemparameter beider nahe beieinander liegen und innerhalb des Konfidenzintervalls bleiben. Bei einigen Items (7, 14, 15, 17, 20) sind die Itemschwierigkeiten der Psychosomatikstichprobe etwas leichter, und bei drei Items (6, 18 und 21) sind sie schwerer (sieht man von Item 9 [Suizidaliät] ab, das generell "schwierig" ist). Obwohl die Abweichungen der Itemparameter damit nicht übermäßig groß erscheinen, weist der Andersen-Test mit LR = 30442.1, df = 62, p < .001 eine signifikante Abweichung nach.

Für die Stichprobe der Gesunden ergibt bereits die visuelle Inspektion der Itemparameter, dass sie in vielen Fällen deutlich abweichend sind. So erscheinen bei den Gesunden mehrere Items, die leichter sind, hauptsächlich aus dem vegetativen Bereich (Schlaf, Appetit, Unruhe und Reizbarkeit). Eher kognitive Items (Schuldgefühle, Entschlusslosigkeit, Wertlosigkeit) sind dagegen schwieriger für die Gesunden; dazu nicht ganz passend: Energieverlust. Bei der Interpretation der Abstände zwischen den Itemparametern muss beachtet werden, dass die Summe der Itemparameter in jeder Stichprobe auf Null normiert ist und deshalb einzelne Items wie v. a. Suizidalität, die wegen extremer Schwellenwerte nur ungenau zu schätzen sind, entsprechend "ausgeglichen" werden müssen. Aber auch die formale Prüfung mit dem Andersen-Test ergibt, dass die Gesunden von den Patienten (beide Gruppen zusammengefasst) signifikant abweichen (LR = 18881.8, df = 62, p < .001).

Stichprobenübergreifend gute Items scheinen zu sein: Traurigkeit, Pessimismus, Versagensgefühle, Selbstkritik, Interessenverlust, Konzentrationsschwierigkeiten. Nach der grafischen Inspektion gehörte zu dieser Liste auch Weinen (Item 10), das sich jedoch in allen psychometrischen Analysen als problematisch erweist, v. a. wegen Kategorie 3 (siehe auch die Schwellenparameter in Abb. 1).

Rasch-Modell (RM)	Anzahl Parameter	BIC	CAIC	AIC	Reliabilität	Klassen- größe
Depressive Patienten						
1-RM	47	4447.2	4494.2	4278.8	.84	100
2-RM	93	4585.4	4678.4	4252.1	.73/.88	53/47%
1-RM smoothed	25	4352.8	4377.8	4263.2	.84	100
2-RM smoothed	51	4392.8	4443.8	4210.0	.74/.92	62/38%
Psychosomatik-Patienten						
1-RM	47	15025.2	15072.1	14799.6	.83	100
2-RM	93	15058.3	15151.3	14611.9	.86/.81	51/49%
1-RM smoothed	25	14906.1	14931.1	14786.1	.83	100
2-RM smoothed	51	14852.4	14876.4	14580.6	.81/.86	53/47%
3-RM smoothed	77	14895.1	14972.1	14525.4	.86/.81/.86	45/44/11%

Tabelle 4: Vergleich der Modellanpassung (PCM) für die Subskala kognitiv in den beiden Patientenstichproben

Die drei Items der Kurzskala (4, 12, 13) zählen nicht zu den durchgängig nahe beieinander liegenden, denn die Gesunden haben deutlich höhere Itemparameter bei Item 4 (Verlust an Freude) und besonders bei Item 13 (Entschlusslosigkeit).

Insgesamt ergeben sich für einige Items gute Übereinstimmungen, aber es zeigt sich doch, dass die Gesunden einige Items anders zu verstehen scheinen als die Probanden der beiden anderen Stichproben, die in sich ähnlicher sind. Ein formaler Test über alle drei Stichproben (Anderson-Test) ergibt, wie schon anhand der Einzelvergleiche zu erwarten, eine Unterschiedlichkeit aller drei Gruppen (LR = 19380.9, df = 124, p < .001).

## 3.3 Die Subskala kognitiv

Die Modellvergleiche für diese Subskala ergeben ein analoges Bild zu den Ergebnissen für die Gesamtskala. Für die Stichprobe Depressiver wird jeweils eine 1-Klassen-Lösung favorisiert, wobei sich unter der Annahme einer geglätteten Verteilung wieder eine Verbesserung ergibt. Für die Psychosomatikpatienten wird aber jeweils eine Klasse weniger empfohlen als bei der Gesamtskala. Auch für diese Stichprobe kann daher Rasch-Homogenität angenommen werden; nur bei smoothed wird (vermutlich bedingt durch die Ersparnis der Parameter) eine 2-Klassen-Lösung vorgeschlagen.

### 4 Diskussion

Die Überprüfung des BDI-II an drei verschiedenen Stichproben mittels Methoden der probabilistischen Testtheorie lässt sich zusammenfassend so interpretieren, dass die Modellannahmen des Rasch-Modells erfüllt zu sein scheinen für die Stichproben der Depressiven und der Gesunden. Hinweise auf Heterogenität ergeben sich für die Patienten einer psychosomatischen Klinik. Diese Stichprobe ist aufgrund der unterschiedlichen Patientenzusammensetzung erwartbar inhomogener als die beiden anderen. Eine Inspektion der Schwellen- und Personenparameter verweist auf response sets, nämlich auf eine Gruppe von Personen mit Tendenz zur Mitte und eine Gruppe von "Extremankreuzern". Diese beiden response sets wurden bereits in Persönlichkeitstests gefunden (Rost et al. 1999); auch Gollwitzer et al. (2005) konnten mit dem Mixed-Rasch-Modell response sets in einem Ärgerfragebogen identifizieren.

Die Analyse der Ordinalität der Antwortkategorien zeigt, dass die Kategorie 3 in wenigen Fällen (v.a. in den Items 6 und 10) nicht optimal formuliert zu sein scheint und dass in einigen weiteren Items die beiden oberen Kategorien eng beieinander liegen. Bezogen auf die "Breite" der Antwortkategorien ergibt sich für alle drei Stichproben, dass das Partial-Credit-Modell am besten geeignet ist, obwohl es die meisten Parameter benötigt. Die Antwortkategorien sind daher symptomspezifisch unterschiedlich weit auseinander (gemessen an ihren Schwellenparametern) und

132 Ferdinand Keller

bestätigen die für jedes Symptom unterschiedlichen Schweregradformulierungen im BDI-II.

Die spezifische Objektivität (oft als Stichprobenunabhängigkeit bezeichnet) liegt nur teilweise vor. Viele Items weisen zwar ähnliche Itemparameter in allen drei Stichproben auf, einige iedoch sind deutlich abweichend und scheinen vor allem für Gesunde eine andere Bedeutung zu haben. Einerseits verweisen auch die Andersen-Tests auf unterschiedliche Itemparameter in den verschiedenen Stichproben, so dass eine direkte Vergleichbarkeit der BDI-II-Items zwischen den Stichproben nicht vollständig gewährleistet ist. Andererseits ist der Andersen-Test sehr sensitiv gegenüber Verletzungen der Grundannahme konstanter Itemparameter und bei großen Versuchspersonenzahlen führt dies (wie bei den meisten likelihoodbasierten Modellgeltungstests) in der Regel zur Verwerfung der Gleichheitsannahme (vgl. Rost 1999, 2004). Als ergänzende Überlegung schlagen daher Kubinger & Draxler (2007) vor zu fragen, ein wie großer Unterschied zwischen Itemparameterschätzungen inhaltlich derart relevant ist, dass er mit dem statistischen Test entdeckt werden kann und wie groß dazu der Stichprobenumfang sein muss; im Umkehrschluss informiert dieser Ansatz dann auch über für die Praxis irrelevante Modellabweichungen. Entsprechende Teststärkeanalysen stehen jedoch erst für einige Spezialfälle zur Verfügung und gehen über eine grobe Orientierung nicht hinaus (Kubinger & Draxler 2007). Zusammenfassend erscheint daher die weitere Analyse dieser Unterschiedlichkeit, eventuell zusätzlich gesondert nach Geschlecht und Altersgruppe, und der response sets lohnenswert, schmälert aber nicht die insgesamt doch guten psychometrischen Eigenschaften des BDI-II, die sich auch bei den Analysen mit probabilistischen Testmodellen herausarbeiten lassen.

# Qualitative und quantitative Methoden. Eine integrative Sichtweise

Jürgen Rost

Der Begriff der qualitativen Forschung stellt von seiner historischen Entwicklung her kein einheitliches Forschungsparadigma dar (vgl. Rost 1998). Er ist vielmehr aus der Kritik an unterschiedlichen Aspekten der konventionellen empirischen Forschung hervorgegangen und definiert sich daher im Wesentlichen negativ, d. h. dadurch, dass qualitative Forschung eben *nicht quantitativ* ist. Dies bringt uns einer Begriffsbestimmung aber noch nicht näher, denn was sind die kritikwürdigen Merkmale herkömmlicher, als quantitativ bezeichneter Forschung und gibt es überhaupt überzeugende Alternativen hierzu, die sich als Merkmale qualitativer Forschung zusammenfassen ließen?

Die Kritik an herkömmlicher, empirischer Forschung, also an den derzeit anerkannten Forschungsparadigmen, bezieht sich auf so gut wie alle Stadien des Forschungsprozesses. In diesem Beitrag werden sechs Stadien oder Bestandteile des Forschungsprozesses unterschieden, für die diskutiert wird, inwiefern die herkömmlichen Konzeptionen als quantitativ bezeichnet und von alternativen, qualitativen Konzeptionen abgegrenzt werden können. Kempf (2008) beschreibt die Wesensunterschiede qualitativer und quantitativer Methoden durch die Gegenüberstellung zweier typischer Repräsentanten für jedes Paradigma, nämlich der Textinterpretation und Inhaltsanalyse (qualitativ) und der Psychometrie (quantitativ). Im hier vorgestellten Modell der sechs Stadien des Forschungsprozesses wäre die Textinterpretation und Inhaltsanalyse dem Stadium 3 (Erhebung von Daten) und die Psychometrie dem Stadium 4 (Kodierung und Skalierung) zuzuordnen. Kempfs Aussage, dass das Spektrum psychometrischer Modelle eine Brücke von den quantitativen zu den qualitativen Methoden schlägt, wobei die Latent-Class-Analyse ein wichtiges Brückenelement darstellt, ist Anlass für die Hoffnung, dass sich der Graben zwischen den beiden Forschungsparadigmen eher schließen, aber zumindest nicht weiter öffnen wird.

# 1 Die wissenschaftstheoretischen Grundlagen empirischer Forschung

Beginnt man mit der wissenschaftstheoretischen Position, die der Forschung zugrunde liegt, so wendet sich die Kritik zunächst gegen das poppersche Falsifikationsprinzip. Nach diesem Paradigma besteht Forschung darin, aus Theorien abge-

134 Jürgen Rost

leitete Forschungshypothesen wiederholt empirischen Falsifikationsversuchen auszusetzen, bis die Empirie den Forscher dazu zwingt, seine Hypothese fallen zu lassen bzw. zu modifizieren und erneuten Falsifikationsversuchen auszusetzen (vgl. Kriz et al. 1987).

Dieses Forschungsprinzip kann in der empirischen Forschungspraxis tatsächlich zu kritikwürdigen oder gar kuriosen Vorgehensweisen führen. So legt es den Hauptschwerpunkt der theoretischen Arbeit auf die Phase *vor* der Datenerhebung, in der es darum geht, den theoretischen Hintergrund einer empirischen Untersuchung auf wenige Hypothesen zuzuspitzen, über deren Wahrheitsgehalt dann anhand der Daten entschieden werden kann. Die Auswertung der empirischen Daten besteht somit "lediglich" darin, darüber zu informieren, ob die jeweilige Hypothese abgelehnt werden muss oder nicht. Aus den empirischen Daten wird lediglich "Ein-Bit-Information "herausdestilliert (Falsifikation der Hypothese: ja oder nein), während die schwere Aufgabe der Theorien*bildung vor* der empirischen Datenerhebung ausgeführt und abgeschlossen sein muss.

Der "laienhaften" Vorstellung, dass man in der Forschung auf der Basis empirischer Beobachtungen Theorien *bildet*, steht im kritischen Rationalismus eine sehr reduzierte Funktion von Empirie gegenüber: Empirie gibt Auskunft darüber, ob eine erneute Theorienbildung oder -modifikation notwendig ist oder nicht, aber sie gibt keine Auskunft darüber, *wie* dies zu geschehen hat (Holzkamp 1972a). Diese aus der Sicht der qualitativen Forschung fragwürdige Reduktion empirischer Erkenntnis kann insofern mit dem Begriff quantitativ charakterisiert werden, als es letztlich ein quantitatives Maß ist, d. h. eine Mittelwertsdifferenz, ein Korrelationskoeffizient oder ein Likelihoodquotient, das mittels einer zugeordneten Wahrscheinlichkeitsverteilung über die Beibehaltung oder Ablehnung der Forschungshypothese entscheidet. Die Forderung qualitativer Forschung ist es, aus empirischen Daten mehr qualitative Information zu ziehen, also darüber, *wie* die Theorie aussehen könnte, und nicht nur die quantitative Information, inwieweit die Theorie falsifiziert werden muss (Flick et al. 1991, Lamnek 1989, Mayring 1996).

Wie ist diese Kritik zu bewerten? Es kann wohl nicht darum gehen, Popper in seinen grundlegenden erkenntnistheoretischen Aussagen und in der Stringenz seiner Argumentation zu kritisieren. Es ist zweifellos richtig, dass jede theoretische Erkenntnis aus dem Kopf heraus kommen muss (daher Rationalismus) und nicht aus den Daten abgelesen werden kann, wie der Empirismus dies so gerne hätte. Andererseits soll sich der Verstand (die Ratio) nicht irgendwelche Theorien ausdenken, sondern empirische Beobachtungen nutzen, um möglichst sinnvolle Theorien zu bilden. Genau dieser Prozess, wie man Theorien aufgrund von empirischen Daten bildet, um sie später mit Hilfe des Falsifikationsprinzips überprüfen zu können, wird von Popper nicht thematisiert, stellt aber einen wichtigen Schritt der Forschungspraxis dar (Popper 1998).

Angesichts der vielen fragwürdigen Theorien, die in den empirischen Sozialwissenschaften einem Falsifikationsversuch ausgesetzt werden, kommt der Frage, *wie* der Forscher seine Theorie gebildet hat, eine immer größere Bedeutung zu. Belie-

bige, in ihrer Genese nicht begründete Theorien und Hypothesen immer erneuten Falsifikationsversuchen auszusetzen, kann zu einer absurden Landschaft von Forschungsergebnissen führen, die viele Kritiker in der Psychologie oder Pädagogik bereits als gegeben ansehen.

Erkenntnisfortschritt ist ein Wechselspiel von Theorie und Empirie. Eine Erkenntnislehre, die lediglich den *deduktiven* Prozess der Ableitung empirisch prüfbarer Aussagen aus Theorien beschreibt, aber nicht den *induktiven* Prozess der Entwicklung von Theorien auf der Grundlage empirischer Beobachtungen, ist tatsächlich defizitär.

Der Erkenntnisfortschritt lässt sich mit Hilfe eines Spiralmodells beschrieben, in dem sich Theorienbildung und empirische Beobachtung abwechseln (vgl. Abb. 1).

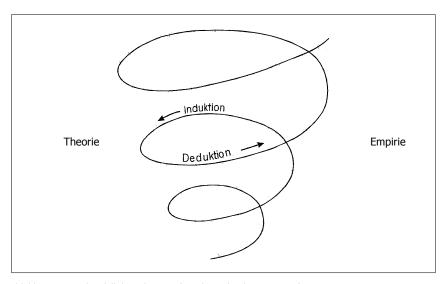


Abbildung 1: Spiralmodell des Erkenntnisfortschritts (nach Sarris 1990)

Aus einer Theorie wird mit Hilfe von deduktiven Prozessen eine Erwartung über empirische Sachverhalte abgeleitet, die einer empirischen Prüfung zugänglich ist. Aus empirischen Beobachtungen wird wiederum durch induktive Prozesse auf die Theorie zurückgeschlossen, diese modifiziert und erweitert. Erkenntnisfortschritt im Sinne einer spiralförmigen Weiterentwicklung kann nur stattfinden, wenn beide Richtungen einer Umdrehung in dieser Spirale funktionieren: Die Überprüfung deduzierter Hypothesen muss ebenso erfolgen wie die Theoriemodifikation infolge der Auswertung empirischer Daten. Qualitative Forschung betont den zweiten Halbkreis, also die Theorienbildung anhand von Daten, und ergänzt somit den deduktiven Halbkreis, der von Popper beschrieben wird. Die Forderung nach qualitativer Forschung bedeutet aus wissenschaftstheoretischer Perspektive, dass Theoriedefizite erkannt und benannt werden müssen, dass Theorien mit Hilfe von

136 Jürgen Rost

empirischen Daten weiterentwickelt werden müssen und Methoden zu entwickeln sind, aus empirischen Daten mittels induktiver Prozesse Theorienbildung zu betreiben.

Dass diesem Theorien generierenden Halbkreis der Spirale des Erkenntnisfortschritts weniger wissenschaftliche Dignität zukommt, liegt – historisch gesehen – am Scheitern des Positivismus. In der Positivismusdebatte zeigte sich, dass induktive Schlüsse von empirischen Beobachtungen auf theoretische Sätze nicht möglich sind, d. h. nicht die Qualität von *logisch wahren* Schlussfolgerungen besitzen (Kutschera 1972). Insofern kann der induktive Weg von empirischen Daten zur Theorie niemals die Qualität logisch wahrer Ableitungen beanspruchen. Andererseits besteht der Wert und Nutzen des menschlichen Denkens vor allem in der Fähigkeit zum induktiven Denken, d. h. zum Ziehen von Schlussfolgerungen aus Wahrnehmungen und Beobachtungen. So gilt auch das induktive Denken (inductive reasoning) als zentrale Fähigkeit der menschlichen Intelligenz (Rost 1997). Wir ziehen ständig (induktive) Schlüsse aus unseren alltäglichen Beobachtungen und tun gut daran, so zu handeln, als ob diese Schlüsse wahr wären, obwohl sie niemals *logisch* wahr sein können.

Wir müssen uns in der empirischen Forschung wohl von der Vorstellung verabschieden, dass wir allein mit logisch korrekten Schlussfolgerungen arbeiten. Wenn wir den Kreis zwischen Empirie und Theorie vollständig durchlaufen wollen (und das müssen wir, um einen Erkenntnisfortschritt im Sinne des Spiralmodells zu erzielen), so müssen wir auch die Induktion von der Empirie auf die Theorie als Bestandteil des wissenschaftlichen Prozesses anerkennen.

## 2 Hypothesen und Forschungsfragen

Am Anfang jeder Forschung steht eine Frage. Diese Feststellung ist ebenso allgemeingültig wie nichtssagend, lässt sich doch jedem wissenschaftlichen oder nichtwissenschaftlichen Erkenntnisvorgang nachträglich eine Fragestellung anheften, z. B. die allgemeine Frage "Wie funktioniert das?" oder "Wie lernen Schüler?".

Was macht eine Frage zum Ausgangspunkt empirischer Forschung bzw. wie muss sie formuliert sein, um als Ausgangspunkt der methodischen Planung einer Forschungsarbeit dienen zu können? In der herkömmlichen Forschung hat man weniger Wert darauf gelegt zu definieren, was eine Fragestellung ist, sondern vielmehr darauf, was eine Hypothese ist. Auch wird nicht eine Fragestellung als Ausgangspunkt betrachtet, sondern vielmehr eine Hypothese. Eine Hypothese stellt eine vorweggenommene Antwort auf eine Frage dar und integriert somit die Fragestellung mit einer antizipierten Antwort. Dass nicht die Fragestellung, sondern eine Hypothese als Ausgangspunkt von Forschung angesehen wird, ist aus der Sicht des kritischen Rationalismus nur konsequent, denn die Empirie kann unsere Fragen nicht beantworten, sondern höchstens Auskunft darüber geben, ob unsere eigenen Antworten (Hypothesen) auf unsere Fragen richtig oder falsch sind.

Was macht man aber nun, wenn man zwar Fragen, aber keine Hypothesen hat? Wie müssen Fragen beschaffen sein, um als Ausgangspunkt von Forschung dienen zu können, ohne dass man bereits Theorien hat, aus denen man erwartete Antworten (Hypothesen) auf diese Fragen ableiten kann? Die Lösung des Problems ist verblüffend einfach: Die Fragen müssen so beschaffen sein, dass sie potenziell durch Hypothesen beantwortbar sind. Unter einer Hypothese versteht man allgemein eine aus einer Theorie abgeleitete Aussage. Definierendes Merkmal einer Aussage ist es, dass sie wahr oder falsch sein kann. Aber was sind das für Aussagen, um die es in empirischer, psychologischer oder pädagogischer Forschung geht? Man spricht hier von Unterschiedshypothesen (z. B. wenn sich Versuchsgruppe und Kontrollgruppe in ihrem Mittelwert voneinander unterscheiden) oder von Zusammenhangshypothesen (z. B. wenn das Merkmal xy eines Schülers mit seinem Lernerfolg zusammenhängt). Ebenso gibt es Entwicklungshypothesen, die zeitliche Verläufe beschreiben, oder prognostische Hypothesen, die das Eintreffen oder Ausbleiben bestimmter Ereignisse vorhersagen, z. B. Berufserfolg.

All diese Hypothesen beschreiben niemals singuläre Ereignisse (z. B. Schüler xy hat das Lernziel der Klasse erreicht), sondern sie beziehen sich auf alle Mitglieder einer bestimmten Stichprobe oder Population. Wenn man einen Lernerfolg der Unterrichtsmethode xy erwartet, so besagt die Hypothese, dass alle Schüler, die diesen Unterricht genossen haben, einen höheren Lernerfolg haben als andere Schüler. Das "alle" meint nicht, dass es für jeden einzelnen Schüler gelten muss, sondern es kann sich auch um eine statistische Aussage handeln, dass eben der Mittelwert dieser Gruppe höher ist. Dies ändert nichts an der Tatsache, dass solche Hypothesen von ihrer logischen Struktur her All-Aussagen sind, d. h. sich immer auf *alle* Mitglieder einer bestimmten Gruppe beziehen.

Betrachtet man die *Art* der Aussagen, die Hypothesen treffen, so kann man feststellen, dass Hypothesen im Bereich psychologisch-pädagogischer Forschung Aussagen über Variablen machen. Eine Variable ist eine Zuordnung von bestimmten Merkmalen zu allen Elementen einer Menge von Objekten, meist Personen. Wenn allen Mitgliedern einer Personenstichprobe das Merkmal männlich oder weiblich zugeordnet wird, so entsteht daraus eine Variable, nämlich die Variable Geschlecht. Ähnlich verhält es sich, wenn man den Personen Messwerte in einem Test zuordnet oder ein qualitatives Urteil über eine erbrachte Leistung. Man kann also sagen, Hypothesen sind Aussagen über eine oder mehrere Variablen, wobei es sich in unserem Bereich meistens um Personenvariablen handelt.

Es lässt sich noch näher spezifizieren, was für Aussagen denn Hypothesen über Variablen machen. Die häufigsten Hypothesen sind, dass zwei Variablen zusammenhängen (korrelieren) oder dass eine Variable in unterschiedlichen Teilstichproben (Versuchsgruppen) unterschiedliche Werte oder Mittelwerte aufweist. Häufig sind auch Aussagen über den Zusammenhang von drei Variablen anzutrefen; z. B. dann, wenn man Interaktionseffekte oder Moderatorwirkungen erwartet (vgl. Kap. 5). Generell gibt es keine Einschränkungen dafür, auf wie viele Variablen sich eine Hypothese bezieht. Gemeinsam haben jedoch all diese Hypothesen, dass sie Aussagen über bestimmte Verteilungseigenschaften von Variablen machen.

138 Jürgen Rost

Eine relativ präzise Definition dessen, was eine Hypothese ist, lautet daher: *Eine Hypothese ist eine Aussage über Verteilungseigenschaften von Variablen* (vgl. Kap. 6).

Nun zurück zur qualitativen Forschung. Wenn man in der qualitativen Forschung schon keine Hypothesen hat, so müssen doch die Fragestellungen so beschaffen sein, dass sie durch Hypothesen, also durch Aussagen über Verteilungseigenschaften von Variablen beantwortbar sind. Dies mag auf den ersten Blick sehr technisch und abstrakt klingen, ist jedoch äußerst hilfreich bei der Ausformulierung von handlungsleitenden Fragestellungen. So macht z. B. die Fragestellung "Wie lernen Schüler im Unterricht xy?" keine Aussage über irgendwelche Variablen und ihre Verteilungseigenschaften, und tatsächlich ist aus einer solchen Frage auch kein sinnvolles methodisches Vorgehen ableitbar. Demgegenüber stellt die Frage "Welche Verständnisschwierigkeiten haben Schüler, wenn sie mit Methode xy unterrichtet werden?" eine Frage über die Verteilungseigenschaften von Variablen. Und zwar wird die Variable Verständnisschwierigkeit angesprochen, die jedem Schüler als Merkmal ein oder mehrere Verständnisprobleme aus einer vielleicht noch unbekannten Palette von solchen Problemen zuordnet. Außerdem gibt es die Variable der Unterrichtsmethode, die jedem Schüler als Merkmal zuweist, nach welcher Methode er unterrichtet wurde.

Die Formulierung von Forschungsfragen als Fragen über die Verteilungseigenschaften von Variablen ermöglicht es, die weitere Planung der Untersuchung mit Hilfe des bestehenden und bewährten Methodeninventars vorzunehmen. So gilt es, im nächsten Schritt für die angesprochenen Variablen zu klären, wie sie definiert sind, wie andere Autoren sie definiert haben, welche Forschungsergebnisse es bislang hierzu gibt, wie man sie operationalisieren kann, welche Probleme es bei ihrer Messung gibt usw. Eine Forschungsarbeit ist um so mehr als qualitativ zu bezeichnen, je weniger gesichertes Wissen es über die beteiligten Variablen gibt. Wurden z. B. die vermuteten Verständnisschwierigkeiten bislang noch nie in dieser Form untersucht, so ist es ein erstes Anliegen der (qualitativen) Forschungsarbeit, zunächst die Existenz der Variable Verständnisschwierigkeiten nachzuweisen. Es ist nämlich durchaus nicht selbstverständlich, dass es gelingt, jedem Schüler genau ein Merkmal aus einer Menge von Verständnisschwierigkeiten zuzuordnen oder auch eine Kombination bestimmter Verständnisschwierigkeiten. Ergeben z. B. Voruntersuchungen, dass jeder Schüler unter einem undefinierbaren Konglomerat verschiedener Verständnisprobleme leidet, so kann man mit dieser (Schein-)Variable keine Forschung betreiben. Aufgabe qualitativer Forschung wäre es dann, unterscheidbare Verständnisschwierigkeiten zu identifizieren, um daraus erst eine Variable konstruieren zu können.

Die Funktion von Forschungsfragen ist es, das weitere methodische Vorgehen zu strukturieren. Forschungsfragen, die als Fragen über Verteilungseigenschaften von Variablen formuliert sind, erfüllen diese Funktion. Inwieweit die zur Beantwortung dieser Forschungsfragen durchgeführte Forschung qualitativ ist, drückt sich darin aus, in welchem Ausmaß neue Variablen erst operationalisiert und identifi-

zierbar gemacht werden müssen und ob über ihre Verteilungseigenschaften bereits Erwartungen vorliegen.

## 3 Die Erhebung von Daten

Zu den von der qualitativen Forschung kritisierten Verfahren der Datenerhebung gehören Fragebögen, standardisierte Testinstrumente, physiologische Messungen, standardisierte Beobachtung und standardisiertes Interview. Demgegenüber gelten narrative Interviews, teilnehmende Beobachtung, Videoaufzeichnungen und die Erhebung verbaler Daten; aber auch Inhaltsanalyse, biografische Daten und ethnografische Formen der Datenerhebung als qualitative Methoden (Kempf 2008, Mayring 1999). Versucht man hierin ein System zu entdecken, so ist es weniger die Unterscheidung von qualitativ und quantitativ, die hier weiterhelfen könnte. Ausschlaggebendes Moment ist vielmehr der Informationsreichtum der Datenquelle. In Multiple-Choice-Fragebögen wird das Verhalten der Menschen auf die Platzierung eines Kreuzes in das richtige Kästchen reduziert. Kennzeichnend für die Fragebogenmethode ist, dass der große Verhaltensspielraum der Menschen in einer Situation einschließlich verbaler, mimischer und gestischer Ausdrucksformen radikal eingegrenzt wird und dass man bei der Datenerhebung auf sehr viel Information verzichtet. Die Reduktion des Verhaltens auf die Platzierung eines Kreuzchens stellt zwar aus der Sicht der "quantitativen" Forscher genau jene Reduktion dar, die für die Beantwortung einer Hypothese benötigt wird. Damit wird aber auf zusätzliche Informationen, wie z. B. die Sicherheit, mit der die Antwort gegeben wird, etwaige Randbedingungen, unter denen die jeweilige Antwort gilt, oder emotionale Begleitreaktionen, verzichtet.

Ähnlich verhält es sich mit Leistungstests, bei denen nur das maximal erreichbare Leistungsniveau erhoben wird, nicht aber die subjektive Anstrengung, die Konzentrationsleistung, die Art der kognitiven Prozesse oder eventuelle Gründe für eine Leistungsminderung oder gar die subjektive Bedeutsamkeit der erbrachten Leistung. Auch hier ist es die Informationsarmut, also die Reduktion auf die Anzahl richtig gelöster Testaufgaben oder die Bearbeitungsgeschwindigkeit, die Tests als quantitative Methoden auszeichnet und von qualitativen Methoden abgrenzt. Ob Interviews oder Beobachtungsmethoden eher den quantitativen oder den qualitativen Verfahren zugerechnet werden, hängt vom Grad ihrer Standardisierung ab. Ein hoher Grad an Standardisierung bedeutet eine Einschränkung des Spektrums beobachtbarer oder erfragbarer Verhaltensaspekte auf jene, die zur Beantwortung der Untersuchungshypothese notwendig sind. Weniger gesteuerte Formen der Beobachtung oder des Interviews, die gar eine Interaktion mit den zu Untersuchenden zulassen, liefern dagegen eine größere Vielfalt an Informationen über die befragten oder beobachteten Personen.

Ganz eklatant wird die Informationsfülle als Definitionsmerkmal qualitativer Methoden, wenn man an Videoaufzeichnungen, Tonbandmitschnitte, umfangreiche verbale Daten, Lebensläufe oder Tagebuchaufzeichnungen denkt. Da der Infor-

140 Jürgen Rost

mationsreichtum das ausschlaggebende Moment qualitativer Datenerhebungsmethoden ist, soll hier auch nicht von quantitativen und qualitativen Daten gesprochen werden (was höchst missverständlich ist), sondern von informationsreichen und informationsarmen Daten(-quellen). Dabei kann 'informationsarm' nicht per se abwertend gemeint sein, denn die Datenreduktion kann sehr sinnvoll auf diejenigen Verhaltensaspekte erfolgen, auf die es in der Untersuchung ankommt. Ebenso ist 'informationsreich' per se noch kein Merkmal hoher Forschungsqualität, denn es kann passieren, dass man nicht weiß, wie man die enorme Informationsfülle reduzieren und für die Theorienbildung nutzen soll.

Dass der Informationsreichtum der Datenquelle dem Anliegen qualitativer Forschung und damit der Kritik an quantitativer Forschung entspricht, ist nicht zufällig: Ist es doch das im vorangehenden Kapitel beschriebene Defizit an einer wissenschaftlich fundierten Theorienbildung aufgrund empirischer Daten, die einen größeren Informationsreichtum empirischer Daten logisch notwendig werden lässt. Es lässt sich sogar die scheinbar paradoxe Beziehung aufstellen: Je weiter fortgeschritten eine Theorienentwicklung ist, desto informationsärmer können die Daten sein, die man zur Überprüfung der Theorie benötigt. Je schwächer und unvollkommener eine Theorie ist, desto informationsreicher sollten die Daten sein, da es nicht nur gilt, die Theorie zu überprüfen, sondern sie zunächst zur Gänze zu entwickeln. Die Kritik an der quantitativen Forschung besteht natürlich nicht darin, dass ihre Theorien schon so weit entwickelt sind, dass man sie mit auf das Wesentliche reduzierten Daten testen kann, sondern der Kern zielt vielmehr darauf ab, dass Theorien, die sich mit informationsarmen Daten testen lassen, eher langweilig, trivial oder vom psychologischen Gehalt her reduktionistisch sind. Man mag der Kritik zustimmen, dass mit Test- und Fragebogendaten oft nur uninteressante Theorien untersucht werden. Jedoch ist der Beweis, dass man mit informationsreicheren Datenquellen zu wesentlich interessanteren und spannenderen Theorien gelangt, auch noch nicht erbracht.

# 4 Kodierung und Skalierung

Die Kritik am konventionellen Paradigma der quantitativen Forschung geht dahin, dass die erhobenen Daten, wie informationsreich auch immer sie sein mögen, nach vorgefertigten Kodierungsregeln in Zahlen transformiert und weiterhin zu quantitativen Messwerten verarbeitet werden. In der quantitativen Forschung müssen in der Tat die Kodierungsregeln prä-experimentell, d. h. vor der Datenerhebung festgelegt sein, da dies dem erkenntnistheoretischen Prinzip entspricht, dass die Datenauswertung allein der Bestätigung oder Ablehnung der Forschungshypothese dient. Würde die Datenkodierung und -auswertung selbst von den Daten abhängig gemacht, also das Kodierungsschema erst anhand der erhobenen Daten entwickelt werden, so wäre die Objektivität der Hypothesenprüfung eingeschränkt. Es wäre möglich, dass die Entwicklung eines Kodiersystems anhand der Daten so erfolgt, dass die Forschungshypothese damit eher bestätigt wird. Andererseits hat es eine gewisse Absurdität, informationsreiche Daten zu erheben,

wenn man sie dann doch auf Auswertungskategorien reduziert, die schon *vor* der Untersuchung feststanden (Strauss & Corbin 1996). Die Erhebung informationsreicher Daten ist überflüssig, wenn man diesen keine Chance lässt, den Forscher von der Unzulänglichkeit seines Kategoriensystems zu überzeugen. Es gibt jedoch Möglichkeiten, ein Kategoriensystem anhand der Daten zu entwickeln und trotzdem der Forderung nach einer objektiven Datenauswertung gerecht zu werden. Es kann z. B. die nach einem Zufallsprinzip aufgeteilte erste Hälfte der Datenmenge zur Entwicklung eines Auswertungssystems benutzt werden, um dieses dann bei der zweiten Hälfte der Daten zur Hypothesentestung einzusetzen. Auch in der qualitativen Forschung kann und sollte die Intersubjektivität der Datenkodierung kontrolliert und überprüft werden. Dies ist z. B. dadurch realisierbar, dass alle Daten nach Entwicklung des Kategoriensystems erneut von unabhängigen Kodierern ausgewertet werden und anhand einer doppelt ausgewerteten Teilmenge der Daten die Übereinstimmung zwischen zwei unabhängigen Kodierern berechnet wird (vgl. Rost 2004).

Sind die Daten kodiert, so sind in einem zweiten Schritt der Datenreduktion Messwerte für die Variablen der Untersuchung abzuleiten. Die Kritik an quantitativer Forschung bedeutet nicht, gänzlich auf Messwerte zu verzichten, denn man benötigt sie, um die Variablen der Untersuchung zu erfassen und die Forschungsfrage zu beantworten. Allerdings gilt auch hier, dass bei herkömmlicher quantitativer Forschung die Messvorschrift prä-experimentell festliegt oder festliegen sollte, während es in der qualitativen Forschung der erste Akt der Datenauswertung ist, geeignete Messwerte zu identifizieren und somit die Variablen zu operationalisieren.

Bei diesem Schritt der Bildung von Messwerten ist erstmals der Unterschied zwischen quantitativer und qualitativer Forschung wörtlich zu nehmen: Üblich ist die Bildung *quantitativer* Messwerte, sei es durch die Bildung von Summenscores oder den Einsatz der Faktorenanalyse, die auf heuristische Weise quantitative latente Variablen ermittelt. Dass es auch möglich ist, *qualitative*, also kategoriale Variablen zu erheben und mit entsprechenden Messvorschriften messbar zu machen, ist nicht nur in der quantitativen, sondern auch in der qualitativen Forschung weitgehend unbekannt (Kempf 2008).

Messen wird in der Regel als synonym zu quantifizieren verstanden; nicht nur in der naturwissenschaftlichen, sondern auch in der sozialwissenschaftlichen Forschung. Dabei wird übersehen, dass die zu messenden Konstrukte in der Sozialwissenschaft oft das Skalenniveau einer Nominalskala haben, also typologische Konstrukte sind. Beispiele sind etwa kognitive Stile, Attributionsstile, kognitive Strategien (Rost 1995), Arten der Zielorientierung im Leistungsverhalten (Köller et al. 1998) oder die Stufen intrinsischer Motivation nach Deci & Ryan (1985). Obwohl solche Konstrukte qualitative Unterschiede zwischen Personen beschreiben, werden sie in der empirischen Forschung in der Regel mit quantitativen Variablen erfasst.

In den letzten Jahren sind Verfahren auf der Grundlage der Latent-Class-Analyse (Lazarsfeld & Henry 1968) entwickelt worden, die es erlauben, kategoriale, also

142 Jürgen Rost

qualitative Personenvariablen zu messen (Rost 2004). Ausgangspunkt dafür sind kategoriale oder ordinale Daten, z. B. Antworten in einem Fragebogen, kategorisierte Beobachtungsdaten oder Interviewdaten. Es werden nicht Summenwerte über die Itemantworten gebildet, sondern die Personen werden danach klassifiziert, wie ähnlich das Profil ihrer Itemantworten ist. Dieses qualitative Konzept von Messen, d. h. die Klassifikation von Personen anhand ihres Antwortprofils, ist in probabilistischen Testmodellen eine interessante Alternative zur Quantifizierung, wie sie in der klassischen Testtheorie und der Item-Response-Theorie angestrebt wird. Es lässt sich bei den latent-class-analytischen Verfahren wie auch z. B. beim Rasch-Modell prüfen, ob die Items tatsächlich die gewünschte quantitative oder kategoriale Variable messen. Speziell ist auch der Vergleich möglich, ob die durch eine Menge von Items erfasste Personenvariable eher quantitativ oder eher kategorial ist. Man braucht für diese Art der Testauswertung allerdings größere Personenstichproben (mindestens 60 bis 70 Personen).

Hinsichtlich der Skalierung der Variablen sollte der Unterschied zwischen qualitativer und quantitativer Forschung nicht darin bestehen, ob eine messtheoretisch fundierte Skalierung vorgenommen wird oder nicht, sondern darin, ob qualitative Personenvariablen tatsächlich als solche skaliert werden. Darüber hinaus kommen Auswertungsverfahren wie die Latent-Class-Analyse auch für eine explorative Skalenanalyse in Frage, wie sie im Rahmen qualitativer Forschung meistens erforderlich ist. Es werden keine prä-experimentellen Hypothesen über die Art der latenten Klassen vorausgesetzt, sondern die Latent-Class-Analyse ermittelt eine Klassifikation der Personen, die im Sinne einer Theorienbildung anhand von Daten interpretiert werden kann.

## 5 Das Forschungsdesign

Das in der klassischen, sprich quantitativen Forschung vorherrschende Design ist sicherlich das einer faktoriellen Versuchsplanung mit einer oder mehreren unabhängigen Variablen und der abhängigen Variablen, auf die eine Wirkung der unabhängigen Variablen erwartet wird. Als quantitative Forschung werden aber auch großflächige Fragebogenuntersuchungen bezeichnet, in denen eine Vielzahl von (quantitativen) Variablen erhoben und miteinander korreliert wird. Die beiden Prototypen experimenteller und nicht-experimenteller empirischer Forschung sind in ihrer Begründung und Aussagekraft derart unterschiedlich, dass man sich fragen kann, inwiefern sie beide in den gemeinsamen Topf der quantitativen Forschung getan werden können.

So benötigt man für experimentelle Studien im Allgemeinen kleinere Stichproben, während für Fragebogenerhebungen sehr große Stichproben erforderlich sind. In experimentellen Studien ist man dazu gezwungen, prä-experimentelle Hypothesen festzulegen. Fragebogenerhebungen werden dagegen allzu oft ohne besonders begründete Theorien durchgeführt. In experimentellen Studien wird versucht, möglichst alle Drittvariablen zu kontrollieren, während Fragebogenerhebungen im

Allgemeinen Feldforschung darstellen und dem weißen Rauschen einer Vielzahl von unkontrollierbaren Drittvariablen ausgesetzt sind. Was haben diese Forschungsarten gemeinsam, was sie aus der Sicht der qualitativen Forschung disqualifiziert?

Gemeinsam ist ihnen z. B. das Missverhältnis von organisatorischem Aufwand für die Datenerhebung (Planung eines Experiments oder Vorgabe eines komplexen Fragebogens an einer großen Stichprobe) und dem Ergebnis, also der Information, ob ein Mittelwertunterschied statistisch bedeutsam ist (im Fall experimenteller Studien) oder ein Korrelationskoeffizient signifikant von Null abweicht (im Fall von Fragebogenerhebungen). Für beide Arten von (quantitativer) Forschung haben sich im Lauf der Zeit gewisse Standards und dazugehöriges Handwerkszeug entwickelt, so dass sich die Bewertung von Forschung oft daran zu orientieren scheint, inwieweit das Handwerkszeug professionell eingesetzt wurde. Welche Erkenntnis für die Weiterentwicklung der Theorien mit dem hohen empirischen Aufwand gewonnen wurde, bleibt oft nebensächlich.

Gibt es denn typische qualitative Forschungsdesigns, die man den beiden klassischen quantitativen Forschungsdesigns gegenüberstellen könnte? Nun, ein qualitatives Forschungsdesign scheint sich durch kleine Stichproben, mögliche Interaktionen zwischen Versuchspersonen und Versuchsleiter, informationsreiche Daten-Aufnahme und minimale Restriktivität der Rahmenbedingungen auszuzeichnen. Mayring (1999) nennt als typisch qualitativ-orientierte Forschungsdesigns die Einzelfallanalyse nach der biografischen Methode, die Dokumentenanalyse, Handlungsforschung, Feldforschung nach der ethnografischen Methode und formative Evaluationsforschung (Kempf 2008). Es stellt sich die Frage, ob im Rahmen derartiger Forschungsdesigns neue Regeln oder Prinzipien der Versuchsplanung gelten oder die klassischen bewährten Prinzipien sinngemäß anzuwenden sind.

An drei zentralen Beispielen soll im Folgenden deutlich gemacht werden, dass die abstrakten Prinzipien der Versuchsplanung (Huber 1987, Kerlinger 1975) auch in der qualitativen Forschung – wenn auch in abgewandelter Ausgestaltung – angewendet werden können und müssen.

Das erste Beispiel ist die Forderung nach einer Kontrollgruppe. Es gibt viele Kritikpunkte am klassischen Kontrollgruppendesign: Die Realisierung der Kontrollbedingung erfordert gleichen Aufwand wie die der Versuchsbedingung, ist aber von
vornherein inhaltlich uninteressant, da die theoretisch interessante Maßnahme in
ihr gar nicht realisiert ist. Sie muss "mitgeschleppt" werden, nur um nachzuweisen,
dass die Versuchsbedingung besser ist. Warum sie nun wirklich besser ist, geht
aus dem Vergleich oft gar nicht hervor. Daraus den Schluss zu ziehen, dass auf
jegliche Vergleichsgruppen verzichtet werden kann, wäre jedoch fatal. Auch in der
qualitativen Forschung ist es sinnvoll, nicht nur eine homogene Gruppe zu untersuchen, da erst der Vergleich mit Gruppen, in denen etwas anderes passiert, Aufschluss darüber gibt, welches Ergebnis spezifisch für die untersuchte Population
ist. Gerade der Prozess der Generierung von Erkenntnis erfordert den Kontrast.
Erkenntnisse darüber, wie z. B. Grundschüler in einem problemlösenden Unterricht

144 Jürgen Rost

lernen, gewinnen erst dadurch an Bedeutung, dass man weiß, unter welchen anderen Bedingungen sie anders lernen würden.

Es gibt mehrere Alternativen zur klassischen Kontrollgruppe, die die Funktion erfüllen, einen Vergleichsmaßstab für die Größe und Bedeutsamkeit von empirischen Befunden zu liefern. Eine mögliche Alternative besteht darin, die Wirkung bestimmter Maßnahmen, z. B. zur Förderung einer autonomen Lernmotivation, in verschiedenen Dosierungen zu untersuchen. Damit ist ebenfalls eine Abschätzung der Effektstärke des untersuchten Unterrichts möglich, ohne eine Kontrollgruppe mit sogenanntem Normalunterricht in die Untersuchung einbeziehen zu müssen. Eine andere Alternative ist es, die interessierende Personenpopulation nicht mit einer "Normalpopulation" als Kontrollgruppe zu vergleichen, sondern mit anderen Populationen, zu denen deutliche Unterschiede erwartet werden. Sollen z. B. besstimmte Leistungsmerkmale von Waldorfschülern untersucht werden, so können anstelle einer schwer definierbaren "Normalstichprobe" auch Schüler von Schulen mit einem ausgeprägten musischen Zweig, Leistungskursschüler mit einem bestimmten Schwerpunkt oder Schüler der Montessoripädagogik herangezogen werden. Solche wohldefinierten Vergleichsgruppen sind aussagekräftiger als eine schlecht definierte Kontrollgruppe, da in ihnen bestimmte Merkmale der zu untersuchenden Maßnahme (Waldorfpädagogik) in unterschiedlicher Kombination und Ausprägung realisiert sein können. Die klassische Kontrollgruppe definiert sich durch die Abwesenheit einer bestimmten Maßnahme, was im pädagogischen Kontext oft nicht sinnvoll zu realisieren ist. Andersartig definierte Vergleichsgruppen können dieselbe Funktion erfüllen, nämlich einen Maßstab zur Beurteilung der Ergebnisse zu liefern, und geben darüber hinaus noch detailliertere Hinweise, worauf die beobachteten Gruppenunterschiede zurückzuführen sein könnten.

Das zweite zentrale Prinzip der Versuchsplanung betrifft die Kontrolle von Störvariablen. Ein Kritikpunkt an quantitativer Forschung besteht darin, dass mit relativ hohem Aufwand Laborbedingungen geschaffen werden, extrem homogene oder parallelisierte Stichproben rekrutiert werden, möglicherweise relevante Drittvariablen ausgeschaltet oder konstant gehalten werden, um in diesem Rahmen einzelne Variablen untersuchen zu können. Dieses Vorgehen schränkt die Aussagekraft der Untersuchung ein, beeinträchtigt ihre ökologische Validität und kann sogar dazu führen, dass das zu untersuchende komplexe Phänomen in dem Versuchsaufbau gar nicht in Erscheinung tritt. Aus dieser Kritik den Schluss zu ziehen, dass man in der qualitativen Forschung gänzlich auf die Kontrolle von Störvariablen verzichten sollte, wäre jedoch fatal. Sowohl in quantitativer wie in qualitativer Forschung geht es darum, auszuschließen, dass die empirisch ermittelten Zusammenhänge zwischen Variablen möglicherweise auf den Einfluss ganz anderer Variablen zurückzuführen sind.

Das zugrundeliegende abstrakte Prinzip heißt Kontrolle von Störvariablen und man mag aus qualitativer Sicht einwenden, dass man noch gar nicht weiß, welche Variable eine Störvariable ist oder ob sie zum untersuchten Phänomen gehört. Dann gilt es, alle in Betracht kommenden Variablen zu benennen und jeweils zu entscheiden, ob sie mit untersucht werden sollen oder nicht. Auch in der qualitativen

Forschung gibt es Störvariablen, die verhindern können, dass sich aus den empirischen Daten diejenigen Erkenntnisse ableiten lassen, die für die Theorienbildung wichtig wären.

Man unterscheidet verschiedene Formen der Kontrolle von Störvariablen, nämlich das Konstanthalten, die Randomisierung, Messwiederholungen, die Einbeziehung in das Forschungsdesign und die statistische Kontrolle. Keine dieser Kontrolltechniken ist von vornherein für qualitative Forschung untauglich, jedoch sind bestimmte Techniken, z. B. die statistische Kontrolle von Störvariablen, eher dem Sachverhalt angemessen, dass in qualitativen Forschungsdesigns oft nur schwache Annahmen über potenzielle Störvariablen existieren.

Das dritte zentrale Prinzip der Versuchsplanung besteht darin, mögliche *Moderatorvariablen* zu entdecken und das Ausmaß ihres Einflusses abzuschätzen. Worin unterscheiden sich Störvariablen, die es zu kontrollieren gilt, von Moderatorvariablen? Störvariablen korrelieren mit Variablen, an deren Zusammenhang man interessiert ist. Korreliert z. B. eine Variable A mit der theoretisch interessierenden Variable B, aber auch mit der potenziellen Störvariable C, so besteht eine Unsicherheit, ob sich tatsächlich A und B gegenseitig beeinflussen oder ob die Störvariable C für den beobachteten Zusammenhang verantwortlich ist. Eine Moderatorvariable korreliert dagegen nicht notwendigerweise mit A oder B, sondern sie verändert (moderiert) den Zusammenhang zwischen A und B, d. h. dieser ist unterschiedlich für verschiedene Ausprägungen der Moderatorvariablen C.

Die Suche nach Moderatorvariablen dient der Absicherung von Kausalinterpretationen, denn ein empirisch gefundener Zusammenhang zwischen zwei Variablen oder ein Mittelwertunterschied zwischen zwei Gruppen sagt nur dann etwas aus, wenn dieser Zusammenhang oder Unterschied invariant ist für alle denkbaren Untergruppen. Hat z. B. ein bestimmter Unterricht eine nachweisbare Wirkung auf die Lernmotivation der Schüler, so kann die beobachtete Motivationsänderung nur dann auf den Unterricht zurückgeführt werden, wenn dieser Effekt bei Jungen und Mädchen, bei älteren und jüngeren Schülern, in verschiedenen Fächern etc. gleich ist. Ist dies nicht der Fall, so muss das Ergebnis: "Der Unterricht hat einen Effekt auf die Lernmotivation" eingeschränkt oder modifiziert werden, etwa derart, dass der Unterricht besonders bei leistungsschwachen Jungen der 9. Jahrgangsstufe einen Effekt auf die Lernmotivation hat.

Auch kann sich bei der Suche nach Moderatoren ein bedeutsamer Effekt in bestimmten Teilgruppen zeigen, der in der Gesamtstichprobe nicht beobachtbar war. Z. B. kann sich ein positiver Effekt bei leistungsschwachen Schülern mit einem negativen bei leistungsstarken Schülern gegenseitig neutralisieren, so dass man bei einer Analyse der Gesamtpopulation zu dem Schluss kommt, der Unterricht habe *keinen* Effekt (vgl. Kap. 6).

Diese Grundprinzipien der Versuchsplanung, Bildung einer Vergleichsgruppe, Kontrolle von Störvariablen und Identifizierung von Moderatorvariablen, müssen ebenso für die Planung einer qualitativen Studie gelten wie für die einer quantitativen. Es gibt keinen vernünftigen Grund, weswegen man in qualitativen Forschungsde-

146 Jürgen Rost

signs darauf verzichten könnte. Ganz im Gegenteil: Wenn deren Ziel darin besteht, die Brücke von den empirischen Daten zurück zur Theorie zu schlagen (vgl. Kap. 1), so sind diese Prinzipien der Versuchsplanung umso bedeutsamer, denn sie tragen dazu bei, eine theoretische Interpretation empirischer Beobachtungen abzusichern.

### 6 Die Auswertung der Daten

Qualitative Forschung beantwortet Fragen, die sich auf die Verteilungseigenschaften von Variablen beziehen. Dementsprechend lassen sich Auswertungsverfahren danach einteilen, ob sie sich auf die Verteilungseigenschaften einer Variablen oder von zwei, drei oder mehreren Variablen beziehen. Oft geht es dabei um den Nachweis und die Exploration einer einzelnen Variable, z. B. die Unterscheidung verschiedener Lernmotivationen beim Schüler oder die Erhebung vorunterrichtlicher Schülervorstellungen zu einem bestimmten Sachgebiet. Unabhängig davon, ob es sich hierbei um eine quantitative Variable (Ausmaß der Lernmotivation) oder um eine qualitative (Art der Lernmotivation) handelt, besteht der erste Auswertungsschritt darin, die Varianz der Variable zu berechnen und zu untersuchen. Nicht zufällig haben die Begriffe Variable und Varianz denselben Wortstamm, d. h. eine Variable ohne Varianz ist keine Variable, sondern eine Konstante. So banal diese Feststellung auch sein mag, sie betont die Notwendigkeit, dass jede Variable im Untersuchungsdesign eine möglichst große Varianz haben muss. Nur was variiert, ist einer empirischen Analyse zugänglich; haben z. B. alle Schüler dieselbe Lernmotivation, so kann auch durch qualitative Forschung nichts über die Zusammenhänge dieser Variable mit anderen Variablen herausgefunden werden.

Die Varianz ist neben dem Mittelwert die zweite zentrale Eigenschaft einer Verteilung. Beide Verteilungseigenschaften müssen gemeinsam interpretiert werden, denn ein Mittelwert mit geringer Varianz sagt etwas anderes aus als ein Mittelwert mit hoher Varianz. Beide Verteilungsmerkmale sind wiederum in Abhängigkeit von der *Form* der Verteilung zu interpretieren: Ist die Verteilung einer Variable zweigipflig (bimodal), so ist der Mittelwert eine missverständliche Größe und die relativ große Varianz nur Ausdruck der Bimodalität.

Der nächste Schritt der Datenauswertung besteht darin abzuschätzen, welcher Anteil der beobachteten Varianz auf den (stets gegebenen) Messfehler zurückzuführen ist und welchen Anteil die wahre Varianz hat.

Den Anteil der wahren Varianz an der beobachteten Varianz einer Variable bezeichnet man als Reliabilität, die auf verschiedene Weise berechnet oder abgeschätzt werden kann (Lienert & Raatz 1994). Die Berechnung der Reliabilität einer Variable ist dann wichtig, wenn es um die Identifizierung und den Nachweis neuer, theoretisch interessanter Variablen geht.

Viele Fragestellungen qualitativer Forschung beziehen sich jedoch nicht auf die Verteilungseigenschaften einer Variable, sondern auf zwei Variablen. Man hat es mit einer bivariaten Verteilung zu tun, zum Beispiel der gemeinsamen Verteilung von Lernmotivation und Lernerfolg. Sind beide Variablen dieser bivariaten Verteilung quantitativ, so ist die interessanteste und am häufigsten berechnete Verteilungseigenschaft die Korrelation. Der Korrelationskoeffizient drückt aus, inwieweit beide Variablen gleichsinnig variieren, d. h. kovariieren. Geht hohe Lernmotivation mit hohem Lernerfolg einher, so ist die Korrelation positiv, geht sie mit einem geringen Lernerfolg einher, so wäre die Korrelation negativ.

Ist eine Variable quantitativ und die andere kategorial, z. B. bei einem Gruppenvergleich, so ist der Mittelwertunterschied die zentrale Eigenschaft dieser bivariaten Verteilung. Der Mittelwertunterschied der quantitativen Variable in den zwei oder mehr Versuchsgruppen drückt den Zusammenhang zwischen unabhängiger und abhängiger Variable aus, ist also ein Zusammenhangsmaß. Je größer der Mittelwertunterschied, desto höher ist die Korrelation.

Sind beide Variablen kategorial, z. B. die Art der Lernmotivation und die Art der Aufgabenlösung, so repräsentiert eine Kreuztabelle die bivariate Verteilung beider. Der Zusammenhang beider Variablen drückt sich in einer Kreuztabelle darin aus, dass einige Zellen häufiger besetzt sind, als es aufgrund der beiden (univariaten) Verteilungen und der Annahme, dass diese beiden Variablen unabhängig sind, zu erwarten wäre. Die Stärke des Zusammenhangs lässt sich in einem Kontingenzmaß ausdrücken.

Der Zusammenhang zweier Variablen, sei es eine Korrelation, eine Kontingenz oder ein Mittelwertunterschied, ist sicherlich die interessanteste Verteilungseigenschaft einer bivariaten Verteilung. Jedoch repräsentieren diese Zusammenhangsparameter noch nicht die Art oder Qualität des Zusammenhangs. So kann der Zusammenhang zwischen zwei quantitativen Variablen nicht-linear sein oder die Kontingenz zwischen zwei kategorialen Variablen nur auf dem Unterschied zweier Valenzen der einen Variablen beruhen. Hierauf kann jedoch nicht weiter eingegangen werden (dazu z. B. Bortz 1984).

Viel interessanter wird die Angelegenheit, wenn drei Variablen an einer Fragestellung beteiligt sind. Es gibt sehr viele verschiedene Möglichkeiten, die trivariate Verteilung zu beschreiben, und verschiedene Begriffe kennzeichnen unterschiedliche Arten des Zusammenhangs von drei Variablen.

Die einfachste und am häufigsten angewendete Art der Beschreibung besteht in der Rückführung auf die drei bivariaten Zusammenhänge. D. h., die trivariate Verteilung von den drei Variablen A, B und C wird zu beschreiben versucht, indem man die drei Korrelationen AB, BC und AC betrachtet. Dies ist z. B. dann der Fall, wenn man anhand der Korrelationsmatrix aller Variablen eine Faktorenanalyse oder ein lineares Strukturgleichungsmodell berechnet. Das Bemerkenswerte an diesem Vorgehen ist, dass sich "echte" trivariate Zusammenhänge gar nicht in den bivariaten Zusammenhängen niederschlagen müssen bzw. durch sie sehr reduziert und verzerrt wiedergegeben werden können. Abbildung 2 verdeutlicht dies an einem Beispiel.

148 Jürgen Rost

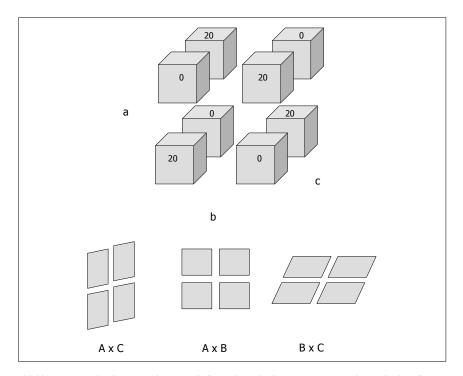


Abbildung 2: Eine dreidimensionale Kreuztabelle, in der jeder bivariate Zusammenhang gleich Null ist

Der in Abbildung 2 dargestellte Würfel zeigt den Zusammenhang von drei dichotomen Variablen in Form einer dreidimensionalen Kreuztabelle sowie die drei bivariaten Kreuztabellen derselben Variablen. Wie man sieht, zeigt sich in keiner der bivariaten Kreuztabellen ein Zusammenhang zwischen den beiden beteiligten Variablen, d. h. die Zellen der Kreuztabellen sind gleich häufig besetzt. Demgegenüber drückt der Würfel aus, dass ein sehr starker trivariater Zusammenhang besteht, dieser ist sogar maximal. Vier der acht Zellen des Würfels sind gleich häufig besetzt, die anderen vier weisen Nullhäufigkeiten auf.

Dieses Beispiel soll illustrieren, wie fahrlässig man handelt, wenn man multivariate Zusammenhänge lediglich über die bivariaten Zusammenhänge analysiert, wie dies in der gesamten Korrelationsstatistik passiert. Komplexe multivariate Zusammenhänge können durch Faktorenanalyse oder Strukturgleichungsmodelle gar nicht abgebildet werden, sondern nur ihre bivariaten Projektionen. Dabei bietet die Methodenlehre eine Reihe von Konzepten an, die echte trivariate Zusammenhänge beschreiben.

Der Begriff der Interaktion ist wohl der geläufigste. Eine Interaktion ist der Zusammenhang zwischen einer Variablen (der abhängigen) von der bivariaten Verteilung zweier anderer Variablen (der unabhängigen). Der Lernerfolg A kann z. B. von der Interaktion der Unterrichtsmethode B und des Geschlechts C abhängen, wenn Mädchen bei der Unterrichtsmethode besser lernen als Jungen.

Während der Begriff der Interaktion symmetrisch ist, beschreibt der Begriff der Moderatorvariablen dasselbe Phänomen aus einer asymmetrischen Perspektive. Das Geschlecht ist diejenige Variable, die den Einfluss der Unterrichtsmethode auf den Lernerfolg moderiert, d. h. verändert. Moderatorvariablen können nur identifiziert werden, indem man die Stichprobe nach den Werten der potenziellen Moderatorvariablen teilt und die Analysen in den Teilstichproben getrennt berechnet.

Auch der Begriff der bedingten Effekte beschreibt eine spezielle trivariate Interaktion. Im genannten Beispiel wäre z. B. der Effekt unter der Bedingung, dass es sich um Mädchen handelt, positiv, während der bedingte Effekt für die Jungen kleiner oder nicht vorhanden ist. Man kann sich vorstellen, dass eine derartige intensive Analyse multivariater Verteilungen umso schwieriger wird, je mehr Variablen an ihr beteiligt sind. Bereits bei vier beteiligten Variablen ist die Anzahl möglicher Interaktionen, Moderatoreffekte oder bedingter Effekte beachtlich groß, so dass man sie in der Regel nur findet, wenn man sie theoriegeleitet sucht. Dies ist ein Punkt, wo das Vorliegen einer Theorie nicht nur ein erkenntnistheoretisches Gebot (s. oben), sondern eine ganz praktische Erfordernis darstellt.

Die Unterscheidung zwischen der Rückführung multivariater Zusammenhänge auf ihre bivariaten Zusammenhänge und der Berücksichtigung "echter" multivariater Zusammenhänge spaltet auch die multivariaten Auswertungsmethoden in zwei Gruppen. Der Großteil der gängigen Statistik beruht auf der Korrelationsstatistik und fokussiert die bivariaten Zusammenhänge einer multivariaten Verteilung. Hierzu gehören, wie bereits erwähnt, die Faktorenanalyse oder auch die Strukturgleichungsmodelle, welche die moderne Form von Pfadmodellen darstellen. Demgegenüber gibt es aber auch Auswertungsmethoden, die die gesamte Information über die multivariaten Zusammenhänge ausschöpfen. Diese Verfahren nennt man Full-Information-Methods. Hierzu gehört z. B. die Latent-Class-Analyse, mittels derer sich die multivariaten Zusammenhänge kategorialer Variablen beschreiben lassen. Die Latent-Class-Analyse sucht nach Mustern von Variablenausprägungen, die besonders häufig auftreten. Sie lässt sich daher gut zur Exploration multivariater Zusammenhänge kategorialer Daten nutzen, ohne dass sie eine Theorie voraussetzt, welche Muster zu erwarten sind.

Dieser Streifzug durch statistische Auswertungsmethoden sollte deutlich machen, dass die bewährten statistischen Verfahren nicht nur eingesetzt werden können, wenn man bestimmte Verteilungseigenschaften erwartet (d. h. Hypothesen hat), sondern auch, wenn man "nur" die Verteilungseigenschaften verschiedener Variablen beschreiben will (d. h. qualitative Fragestellungen hat). Der Unterschied besteht darin, dass man im ersten Fall gezielt bestimmte Verteilungsparameter berechnet, um diese dann mit den Methoden der Inferenzstatistik mit den theoretisch erwarteten Parameterwerten zu vergleichen. Im Fall qualitativer Fragestellungen muss man dagegen sehr viele Verteilungseigenschaften unter die Lupe nehmen, um Phänomene zu entdecken, die zur Theoriebildung beitragen können. Die infe-

150 Jürgen Rost

renzstatistische Absicherung der Ergebnisse erübrigt sich zwar, wenn man keine prä-experimentellen Hypothesen hat, jedoch lässt sich auch hier über die Berechnung von Konfidenzintervallen für die Verteilungsparameter beurteilen, inwieweit beobachtete Phänomene stark genug sind, um interpretationswürdig zu sein (vgl. z. B. Brandstätter 1999).

### 7 Zusammenfassung

In diesem Beitrag wurde eine integrative Sichtweise von gualitativen und guantitativen Methoden dargestellt, die in der Regel mit einem Entweder-oder oder auch einem Sowohl-als-auch argumentierte oder resümierte. Doch es gibt Konzepte der qualitativen Methoden, die hier nicht genannt worden sind, sei es aufgrund fehlender Vertrautheit mit einer Methode oder sei es, dass diese sich nicht in das hier behandelte Schema einpassen lässt. Hierzu seien die Konzepte der sozialpsychologischen Rekonstruktion von Bedeutung genannt, "der genetischen Rekonstruktion von Handlungen" und das gemeinsame Gegenstandsverständnis, "das die soziale Wirklichkeit als das Produkt von Bedeutungszuweisungen begreift, welche ihren Ursprung in der sozialen Interaktion haben und die Grundlage für das Handeln der Menschen bilden" (Kempf 2008, S. 15). Wenn auch die Vertreter der qualitativen Methodenlehre, z.B. mit dem Begriff der Triangulation, ein eigenes Modell für die Koexistenz von qualitativen und quantitativen Methoden entwerfen, scheint es bis heute nicht zu einem wirklichen Konsens über den Stellenwert der verschiedenen Methoden in einem gemeinsamen wissenschaftstheoretischen Rahmen gekommen zu sein.

Vielleicht ist die Suche nach einem derart strengen Konsens aber auch Ausdruck eines übertriebenen Strebens nach sozialer Rekonstruktion von Bedeutungszusammenhängen. Rost (2002) hat bei seiner Suche nach Modeströmungen und Zeitgeist in der Methodenlehre nicht weniger als sieben fundamentale Dichotomien gefunden, deren Präferenz für eine der beiden Alternativen nur mit Zeitgeist und Mode zu erklären ist. Die letzte Dichotomie lautete qualitative vs. quantitative Forschung.

Vorfassung dieses Beitrags in:

Kay Spreckelsen, Kornelia Möller & Andreas Hartinger (Hrsg.) (2002). Ansätze und Methoden empirischer Forschung zum Sachunterricht. Bad Heilbrunn: Klinkhardt.

# Subjektwissenschaftliche Perspektiven und integrative psychologische Forschungsmethodik

# Qualitative versus quantitative Methoden: eine ewig währende Auseinandersetzung?

Elfriede Billmann-Mahecha

"Wir plädieren für einen gegenstandsbezogenen Einsatz verschiedenster Forschungsmethoden, die die klassischen Grenzen von Natur- und Kulturwissenschaften aufhebt und so der Komplexität des Forschungsgebiets Mensch gerecht wird."

Kempf & Kiefer 2009a, S. 10

#### 1 Zwei unvereinbare Paradigmen?

Seit Jahrzehnten gibt es in der akademischen Psychologie den Versuch, qualitative Forschung salonfähig zu machen – mit mäßigem Erfolg. Quantitativ zu forschen, entspricht der Mainstream-Psychologie; mit ihr steht man auf der sicheren Seite, wenn man in anerkannten Journalen publizieren, Drittmittel einwerben und eine gute Position im Bewerbungswettbewerb erreichen will. Quantitative Forschung steht in der Tradition des kritischen Rationalismus, beruht auf logisch (nahezu) einwandfreien Designs, auf messtheoretisch (meist) gut begründbaren Datengenerierungen und auf komplexen statistischen Analyseverfahren. Die quantitative Modellbildung und Methodik hat sich mit bemerkenswerten Fortschritten zu einem respektablen Forschungsfeld entwickelt. Die qualitative Forschung hat sich zwar ebenfalls erheblich weiterentwickelt, wie in einer Fülle an Lehrbüchern und Handbüchern nachzulesen ist (z. B. Mey & Mruck 2010), führt aber nach wie vor eher ein Nischendasein und gilt – bei allem freundlichen Respekt – als "soft" und nicht als "hard" (vgl. Kempf 2003a).

Die wissenschaftstheoretischen und methodologischen Auseinandersetzungen darüber, wie das Psychische und das Soziale, also das komplexe Forschungsgebiet Mensch, angemessen wissenschaftlich zu erforschen sei, haben eine lange Tradition. Man denke an die Versuche, Alternativen zum Hempel-Oppenheim-Schema der Kausalerklärung zu begründen. Entsprechende Überlegungen reichen von Versuchen, das Schema in Form induktiv-statistischer Erklärungen zu retten, über Fragen, ob Handlungserklärungen als Kausalerklärungen konzipiert und verteidigt werden können, bis zu prinzipiellen Alternativen wie etwa der narrativen Erklärung (vgl. Danto 1974, Kochinka & Werbik 1997, Straub 1999).

Bei alledem bleibt die Frage virulent, ob es sich bei qualitativen und quantitativen Zugängen zum menschlichen Erleben, Verhalten und Handeln um zwei miteinander unvereinbare Paradigmen im kuhnschen Sinne handelt, die sowohl in metatheoretischer Hinsicht von unterschiedlichen Menschenbildern (anthropologischen Kernannahmen) ausgehen als auch in wissenschaftstheoretischer Hinsicht in unterschiedlichen Konzeptionen, was Wissenschaft ausmacht und vor anderen Wissensformen auszeichnet, begründet sind. Die Hardliner auf beiden Seiten betonen selbstverständlich die Unvereinbarkeit. Aber es gibt auch Lehrbücher – und damit ist ausdrücklich nicht das dreibändige Werk von Kempf Forschungsmethoden der Psychologie (2003 ff.) gemeint -, die einer Sichtweise der Inkommensurabilität vermutlich ungewollt Vorschub leisten, indem sie aus didaktischen Gründen die Unterschiede zwischen quantitativer und qualitativer Forschung in tabellarischen Übersichten und eingängigen Merksätzen herausstreichen. Seit gut zehn Jahren beginnt sich demgegenüber eine Zwischenposition zu etablieren, die unter dem Label Mixed Methods argumentiert, dass es sich bei qualitativen und quantitativen Ansätzen nicht um kategoriale Unterschiede handelt, vielmehr alle Methoden auf einem Kontinuum zwischen den Polen qualitativ und quantitativ anzusiedeln seien (vgl. Benz & Newman 1998).

Im Folgenden wird zunächst ein kurzer historischer Rückblick auf die Auseinandersetzungen um das wissenschaftliche Selbstverständnis der Psychologie gegeben (Kap. 2; vgl. dazu ausführlicher Billmann-Mahecha 2001), danach erfolgt eine Infragestellung dichotomer Gegenüberstellungen qualitativer und quantitativer Ansätze (Kap. 3), bevor kurz auf Mixed-Methods-Ansätze eingegangen wird (Kap. 4). Abschließend wird der mögliche Gewinn einer Verbindung qualitativer und quantitativer Forschung anhand eines Forschungsbeispiels veranschaulicht (Kap. 5 und 6).

#### 2 Kurzer historischer Rückblick

Kontroversen über anthropologische Grundannahmen und theoretische Richtungen sowie über die Zuordnung zu den Natur- oder Geistes-/Kulturwissenschaften haben die akademische Psychologie schon seit ihren Anfängen begleitet. Solche Kontroversen sind zwangsläufig mit wissenschaftstheoretischen, methodologischen und methodischen Fragen verknüpft. In der Bundesrepublik erreichte die Auseinandersetzung um Methoden einen ersten Höhepunkt in den 1950/60er-Jahren, als sich die damals jüngere Wissenschaftlergeneration um die Rezeption der nordamerikanischen Psychologie bemühte, die während der NS-Zeit von den verbliebenen Fachvertretern so gut wie nicht zur Kenntnis genommen worden war. Der polemisch ausgetragene Streit zwischen der überwiegend ganzheitlich-charakterologisch orientierten deutschen Psychologie und den an einer Reform interessierten Fachvertretern entzündete sich am Streit um die Validität persönlichkeitsdiagnostischer Untersuchungen. Die prominentesten Protagonisten der beiden Lager waren Albert Wellek und Peter Hofstaetter.

Fachvertreter, die sich um eine Integration verschiedener Forschungstraditionen bemüht haben, wie z. B. Hans Thomae, wurden in diesem Streit nicht meinungsführend. Vielmehr stabilisierte sich die naturwissenschaftlich orientierte Ausrichtung, und das geisteswissenschaftliche Erbe geriet zunehmend in Vergessenheit (vgl. Lück et al. 1987) – angesichts mancher nationalsozialistischer Vereinnahmung zum Teil zu Recht. Zum Teil wurden damit aber auch vielversprechende Forschungsansätze aus dem ersten Drittel des 20. Jahrhunderts mit hinweggefegt, die erst viel später aus den USA reimportiert wurden, wie zum Beispiel Ansätze der ökologischen Psychologie (vgl. z. B. die Lebensraumstudie von Martha Muchow aus den frühen 1930er-Jahren [Muchow & Muchow 1998]).

Eine gewisse Erschütterung hat die "neue" naturwissenschaftlich-statistische Orientierung der Psychologie dann ab den 1970er-Jahren erlebt. Der sog. Positivismusstreit in der Psychologie (vgl. Métraux 1985) war einerseits von der Studentenbewegung und andererseits vom Positivismusstreit in der deutschen Soziologie beeinflusst. Die Studentenbewegung hat ihre Psychologiekritik anfänglich allerdings weniger an der Dominanz statistischer Methoden als an Fragen nach der Relevanz psychologischer Forschung festgemacht (vgl. Mattes 1985). In jener Zeit hat auch Klaus Holzkamp (1970) zum ersten Mal seine Kritik an der mangelnden gesellschaftlichen Relevanz psychologischer Forschung vorgetragen. Die Verunsicherung in der akademischen Psychologie aufgrund der Studentenunruhen brachte Graumann 1970 als Vorsitzender der Deutschen Gesellschaft in seinem ersten Bericht zur Lage der Psychologie zum Ausdruck: "Die Fragen, zumindest soweit sie bisher artikuliert und begründet worden sind, betreffen tatsächlich unser wissenschaftliches Selbstverständnis" (Graumann 1973, S. 23).

Die von der Studentenbewegung in erster Linie politisch motivierte Kritik schloss sich im Folgenden an die Position der Kritischen Theorie im Positivismusstreit der deutschen Soziologie an. Die Kritik am einheitswissenschaftlichen Ideal positivistischer Prägung wurde im Laufe der 1970er-Jahre schließlich in breiteren Kreisen der Psychologie reflektiert. Es kam die Rede von der Psychologie als einem pluralistischen System auf (Thomae & Feger 1969, Thomae 1977); andere forderten einen Paradigmenwechsel (Groeben & Scheele 1977); erneut wurde – wie bereits von Karl Bühler 1927 – von einer Krise der Psychologie gesprochen (Mertens & Fuchs 1978), wenn auch mit anderen Konnotationen als zu Bühlers Zeiten.

Im Weiteren wurde von einigen Fachvertreter(inne)n nach Alternativen im Bereich der qualitativen Methoden gesucht (vgl. z. B. den Sammelband von Jüttemann 1985), dies allerdings nicht im Rückgriff auf eigene Forschungstraditionen, sondern im Anschluss an die soziologische Methodendiskussion und -entwicklung. Die zunehmende Verlagerung der Diskussion um den Status der Psychologie als Wissenschaft auf den Gegensatz zwischen qualitativen und quantitativen Methoden führte dazu, dass wieder – wie schon in den 1950er-Jahren, nur mit umgekehrtem Vorzeichen – "die Methodik zum Sündenbock für Fehler gemacht [wird], die auf einer anderen Ebene liegen", wie Bungard (1984, S. 48) die Lage in den 1980er-Jahren beschreibt.

Auffällig an diesen neueren Auseinandersetzungen um die Methodenfrage in der Psychologie ist, dass sich die qualitative Psychologie zunächst in erster Linie über eine Kritik an den vermeintlichen Fehlentwicklungen der quantitativen Psychologie zu legitimieren versuchte, während die etablierte Psychologie mit einem Spektrum reagierte, das von Ignoranz bis zum Vorwurf der bloßen Subjektivität und Unwissenschaftlichkeit reicht. Dabei hält sich bis heute auch hartnäckig das Vorurteil, die qualitative Forschung befasse sich nur mit Einzelfällen. Was Métraux für den Methodenstreit der 1950er-Jahre konstatierte, gilt vergleichsweise auch für den der 1970/80er-Jahre: "Statt ein wohlgeformtes Problem zu analysieren, um es einer zumindest akzeptablen Lösung zuzuführen, uferte der Streit zu einem Zank um die Psychologie und deren Methoden schlechthin aus" (Métraux 1985, S. 240). Auch das Postulat von zwei wissenschaftlichen Psychologien wurde wieder vertreten, so z. B. von Werbik (1985), der dafür die Begriffe Psycho*logie* und Psycho*no*mie vorschlug. In einer späteren Stellungnahme ist von dieser Differenzierung allerdings nicht mehr die Rede: "Der Streit um die Frage, ob die Psychologie eine Natur- oder eine Kulturwissenschaft ist, beruht auf einer falschen Alternative" (Hartmann & Werbik 2001, S. 178). Ähnlich argumentieren heute auch Kempf & Kiefer (2009a).

In der angegriffenen Mainstream-Psychologie jedenfalls stieg in den 1990er-Jahren die Akzeptanz qualitativer Forschungsansätze – nicht zuletzt infolge der langsamen Öffnung amerikanischer Fachzeitschriften für qualitative Arbeiten. In der 1995er Auflage des Standardlehrbuchs *Forschungsmethoden und Evaluation* von Bortz & Döring erschien erstmals ein umfangreicher Abschnitt zu qualitativen Verfahren. Allerdings bleibt ein Problem bestehen: Qualitative Erhebungs- und Auswertungsverfahren beziehen ihre theoretische und methodologische Begründung aus einem weitverzweigten interdisziplinären Diskurs der Sozial- und Kulturwissenschaften, der sich u. a. dem sog. cultural turn verpflichtet sieht. Da dieser Diskurs in den verschiedenen Einzeldisziplinen aber nicht einheitlich interpretiert wird (vgl. z. B. Appelsmeyer & Billmann-Mahecha 2001), ist auch eine einheitliche Forschungslogik der qualitativen Sozialforschung kaum in Sicht. Umso schwieriger ist es, eine Forschungslogik zu entwickeln, die qualitative und quantitative Verfahren nicht nur forschungspraktisch, sondern auch in erkenntnistheoretischer Hinsicht integriert.

### 3 Was taugen dichotome Gegenüberstellungen?

Betrachtet man die Auseinandersetzungen um Methoden aus einer historischen Perspektive, so fällt eine seltsame Dichotomisierung auf, die sich durch die Psychologie des 20. Jahrhunderts zieht und die wir bereits bei Wilhelm Wundt finden: "Demnach verfügt die Psychologie ... über zwei exakte Methoden: die erste, die experimentelle Methode, dient der Analyse der einfacheren psychischen Vorgänge; die zweite, die Beobachtung der allgemeingültigen Geisteserzeugnisse, dient der Untersuchung der höheren psychischen Vorgänge und Entwicklungen" (1911, S. 30). Hat Wundt mit diesem Diktum einer Aufspaltung der Psychologie in zwei

Hauptrichtungen das Wort geredet? Etwas später ist die Rede von: naturwissenschaftlich vs. geisteswissenschaftlich, objektivierend vs. subjektivierend, nomothetisch vs. idiographisch, erklärend vs. verstehend und in jüngerer Zeit qualitativ vs. quantitativ. Jedes dieser Gegensatzpaare hat seine eigenen wissenschaftstheoretischen und -geschichtlichen Hintergründe, die hier nicht ausgeführt werden können. Vergleicht man stattdessen entsprechende Gegenüberstellungen über das Jahrhundert hinweg, dann fallen erstaunliche Parallelen auf. So beginnt z. B. die tabellarische Übersicht, die Müller-Freienfels (1929) Ende der 1920er-Jahre in Die Hauptrichtungen der gegenwärtigen Psychologie publiziert hat, ebenso wie die tabellarische Gegenüberstellung qualitativer und quantitativer Sozialforschung von Lamnek (1995) mit der Gegenüberstellung naturwissenschaftlich vs. geisteswissenschaftlich. Beide heben für die naturwissenschaftliche Richtung u. a. die Bestimmungsstücke Kausalität, Standardisierung und Quantifizierung hervor. Beide erwähnen als charakteristisch für die geisteswissenschaftliche Orientierung die ganzheitliche Betrachtung, die Hermeneutik und die Typenbildung. Eine aktuelle Dichotomisierung findet sich in dem Lehrbuch von Hussy et al. (2010, S. 184), dem die folgenden Zwischenüberschriften entnommen sind.

Zu jedem der tabellarisch aufgeführten Gegensätze kann man wissenschaftstheoretische Argumente und Forschungsbeispiele vorbringen, die eine solche Dichotomisierung ins Wanken bringen. Das soll hier nicht im Einzelnen durchgespielt werden, sondern exemplarisch. Aber fragen wir zunächst: Was sind eigentlich qualitative und was sind quantitative Daten? Diese Frage mag banal klingen und es gibt in der Literatur auch genügend Definitionen. Mit Wörter vs. Zahlen kann man die Unterschiede vielleicht auf den Punkt bringen (Erzberger 1998) oder auch mit verbale Daten vs. numerische Daten. Worauf es im Folgenden ankommt, ist, dass in beiden Fällen das empirisch Vorfindliche, Aufgesuchte oder Hergestellte einer Reihe von Transformationsprozessen unterzogen wird, bis Daten vorliegen, die etwas *über* die Wirklichkeit aussagen sollen. So wird in der qualitativen Forschung z. B. das gesprochene Wort auf einem Tonträger konserviert und mehr oder weniger akribisch transkribiert, d. h. es wird Ort und Zeit enthoben und in einen schriftlichen Text transformiert. Auch visuelles Material - Beobachtungen, Fotos, Videomitschnitte, Zeichnungen etc. – muss in der einen oder anderen Form in verbale Daten, also Wörter, transformiert werden, um es der weiteren Interpretation zugänglich zu machen (vgl. Billmann-Mahecha 2010). Diese mehr oder weniger regelgeleitete, immer aber bereits interpretative Transformation ist im Prinzip nichts anderes, als empirische Sachverhalte in numerische Relationen zu transformieren, wie es für die guantitative Forschung charakteristisch ist.

## 3.1 Naturalistische Vorgehensweise vs. aktive Manipulation?

Vor diesem Hintergrund ist kaum nachvollziehbar, warum qualitative Daten als alltagsnäher oder gar als naturalistisch bezeichnet werden. Auch qualitative Daten können höchst artifiziell sein, man denke nur an die methodologische Diskussion zum narrativen Interview. Es ist für einen Probanden durchaus keine alltagsnahe Situation, einer fremden Person mehrere Stunden lang das eigene Leben zu er-

zählen und davon auch noch eine Bandaufnahme zu gestatten, die möglicherweise von weiteren, unbekannten Personen transkribiert und interpretiert wird.

### 3.2 Interpretationsbedürftige Daten vs. numerische Daten?

Seltsam mutet es auch an, nur qualitative Daten als interpretationsbedürftig zu charakterisieren. Jeder weiß, dass auch quantitative Daten nicht für sich sprechen. Insbesondere komplexe Datensätze können aus verschiedenen theoretischen, methodischen oder auch politischen Perspektiven weiter analysiert und interpretiert werden. Man denke etwa an die (zunächst) kontroverse Interpretation ein und desselben Datensatzes der Berliner ELEMENT-Studie durch Rainer Lehmann und Jürgen Baumert (vgl. Baumert et al. 2009).

Aber was heißt jeweils Interpretation? Haben qualitative und quantitative Forscher(innen) unterschiedliche Interpretationsbegriffe, unterscheidet sich der Stellenwert von Interpretation im Forschungsprozess oder haben qualitative und quantitative Forscher(innen) eine unterschiedliche Sensibilität für die interpretativen Anteile in ihren Arbeiten ausgebildet? In der qualitativen Forschung jedenfalls besteht – in der Regel – eine ausgeprägte Sensibilität dafür, dass jeder Schritt von den vorbereitenden Arbeiten bis hin zur Konstruktion von Typiken, wenn man solche anstrebt, im Forschungsprozess interpretationsbeladen ist und im Zweifelsfall entsprechend gerechtfertigt werden kann und muss (vgl. ausführlich Straub 1999). Aber auch in der quantitativen Forschung ist Interpretationsarbeit von Anfang an erforderlich; sie steckt bereits in der mehr oder weniger theoriegeleiteten Operationalisierung der zentralen Begriffe sowie in der Konstruktion der Erhebungsinstrumente. Interpretation (Sinnexplikation), Abstraktion und Theoriearbeit können in gualitativen und guantitativen Designs zwar an unterschiedlichen Stellen im Forschungsprozess besonders hervorgehoben sein, verzichtbar sind sie aber in keinem Forschungsansatz und vom Grundsatz her zumindest "familienähnlich".

### 3.3 Induktives Vorgehen vs. deduktives Vorgehen?

Hierzu findet sich im Lehrbuch von Hussy et al. der hervorgehobene Merksatz "Qualitative Forschung ist *induktiv*, d. h., sie geht nicht von theoretischen Annahmen aus" (2010, S. 185, Hervorh. im Orig.). Dass qualitative Forschung nicht von theoretischen Annahmen ausgeht, ist ein absoluter Spezialfall, dessen vermutlich gar nicht haltbare Begründung auf die Anfänge der Diskussion um die Grounded Theory zurückgeht. Aber davon abgesehen: Betrachten wir das idealtypische, theorieprüfende Ablaufmodell des quantitativ orientierten Forschungsprozesses, das ein deduktives Vorgehen präskribiert, und rekonstruieren wir konkrete Forschungsabläufe vor Ort, so müssen wir konstatieren, dass die Forschungspraxis meist anders aussieht als in den Journal-Artikeln publiziert: Über nicht haltbare Hypothesen wird nicht berichtet, dafür werden aus dem Datensatz "emergierte" Hypothesen nachträglich an den Anfang gestellt, um nur ein den "Widrigkeiten" der tatsächlichen Forschungspraxis entnommenes Beispiel zu erwähnen. Explizit in-

duktiv verfahren quantitativ orientierte Wissenschaften, wenn sie hochkomplexe Datensätze explorieren, daraus Hypothesen ableiten und diese mittels verschiedener Modellrechnungen überprüfen (vgl. z. B. die molekularbiologische Erforschung von Krankheitsbildern, die Klimaforschung, aber auch die Folgestudien auf der Basis der PISA-Datensätze). Schließlich ist in Bezug auf die Forschungslogik zu konstatieren, dass auf der Wahrscheinlichkeitstheorie beruhende, statistisch gewonnene Allgemeinaussagen letztlich *auch* einer induktiven Forschungslogik folgen (zu deduktiv-nomologischen vs. induktiv-statistischen Erklärungen vgl. ausführlich Kempf 2003a, Kap. 2).

#### 4 Mixed Methods

Über das Verhältnis zwischen qualitativer und quantitativer Forschung und mögliche Verbindungen wird nicht erst nachgedacht, seit der Begriff Mixed Methods in Mode kam (Teddlie & Tasakkori 2009; für einen Überblick vgl. Schreier & Odağ 2010). Bereits von Eye (1994) diskutierte sechs Typen möglicher Beziehungen zwischen qualitativer und quantitativer Forschung: (1) die beiden Ansätze sind antithetisch, (2) sie stehen in einem Komplementärverhältnis zueinander, (3) sie operieren parallel zueinander, (4) sie sind ineinander eingebettet, (5) sie erlauben es, die Methoden des jeweils anderen Ansatzes zu nutzen, (6) die qualitative Forschung geht der quantitativen voraus. Während der erste Typ eine starre, abgrenzende Haltung repräsentiert und insofern als überholt angesehen werden sollte, stellen die Typen 2 bis 6 verschiedene Möglichkeiten dar, wie qualitative und quantitative Forschungsansätze in komplexen Foschungsdesigns miteinander verbunden bzw. aufeinander bezogen werden können (vgl. hierzu auch Mayring 2001). Genau das ist das Anliegen von Mixed Methods – einem Ansatz, der seit gut einem Jahrzehnt vor allem in den USA diskutiert wird.

Nach Schreier & Odağ (2010) bezeichnet Mixed Methods "im weitesten Sinne die Kombination von Elementen eines qualitativen und eines quantitativen Forschungsansatzes innerhalb einer Untersuchung oder mehrerer aufeinander bezogener Untersuchungen" (S. 263). Die Idee ist nicht neu; bereits aus den 1930er-Jahren haben wir mit der Marienthal-Studie (Jahoda et al. 1933; vgl. auch die Verfilmung "Einstweilen wird es Mittag" von Karin Brandauer [1988]) ein prominentes, immer wieder zitiertes Beispiel, wie qualitative und quantitative Methoden ineinander greifen können, um einen interessierenden Phänomenbereich mehrperspektivisch zu untersuchen. Auch Wilhelm Kempf ist die Integration qualitativer und quantitativer Methoden ein wichtiges Anliegen. Mit der latenten Stilanalyse hat er ein vielfach einsetzbares Verfahren entwickelt, mit dem auf der Basis der Latent-Class-Analyse quantitative und qualitative Inhaltsanalysen am selben empirischen Material systematisch aufeinander bezogen werden können. Ein Anwendungsbeispiel findet sich u. a. in Kempf (2002) zur internationalen Kriegsberichterstattung über den Golfkrieg.

Wesentlich älter als die Mixed-Methods-Diskussion ist die um das Konzept der Triangulation, eine Diskussion, die auf Campbell & Fiske (1959) zurückgeht und später in der qualitativen Forschung v. a. von Denzin (1970) aufgegriffen und weiterentwickelt worden ist. Beim Konzept der Triangulation ist auf Methodenebene nicht festgelegt, ob die verschiedenen Methoden aus demselben oder aus verschiedenen Forschungsparadigmen kommen, während beim Mixed-Methods-Konzept prinzipiell qualitative und quantitative Methoden kombiniert werden. Folgt man Flick (2010), so kann die Mixed-Methods-Diskussion – salopp formuliert – der Triangulationsdiskussion nicht das Wasser reichen, weil sie in wissenschaftstheoretischer und methodologischer Hinsicht einen eher pragmatischen Standpunkt einnimmt. Tatsächlich scheint die Frage nach der epistemologischen Begründung im Vergleich zur Entwicklung immer komplexerer Designs in der Diskussion um Mixed Methods bis heute eher nachrangig behandelt worden zu sein.

#### 5 Ein Forschungsbeispiel

Die Typologie von von Eye aufgreifend, wird im Folgenden auf der Basis einer qualitativen Pilotstudie zum mathematischen Selbstkonzept von Mädchen und Jungen einer zweiten Grundschulklasse exemplarisch dargestellt, inwiefern qualitative und quantitative Methoden *komplementär* zueinander eingesetzt werden könnten. Bei dieser Position wird davon ausgegangen, "daß beide Ansätze potentiell wichtige Information zu einer Untersuchung beitragen können, und daß die beiden Informationsquellen sich gegenseitig ergänzen" (von Eye 1994, S. 36).

Ein inzwischen mehrfach bestätigtes Ergebnis der Selbstkonzeptforschung besteht darin, dass bereits in dritten und vierten Grundschulklassen Mädchen über alle Leistungsstufen hinweg ein durchschnittlich geringeres mathematisches Selbstkonzept aufweisen (vgl. Tiedemann & Faber 1995). Vor dem Hintergrund dieser Befunde, die durchweg mit Hilfe quantitativer Verfahren gewonnen wurden, untersuchten wir das mathematische Selbstkonzept von Mädchen und Jungen in einer zweiten Grundschulklasse mit Hilfe des qualitativen Verfahrens der Gruppendiskussion. Ergänzend wurde in einem offenen Interview die Sichtweise der Lehrerin zu den Ausführungen der Kinder erhoben.

An der Pilotstudie haben insgesamt 13 Kinder einer zweiten Grundschulklasse teilgenommen: neun Mädchen in zwei Gruppen und vier Jungen in einer Gruppe. Die Gesprächsleiterin eröffnete die Diskussionen mit der Eingangsfrage "Macht Euch Mathe denn Spaß?" und orientierte sich mit ihren Nachfragen zum einen an den von den Kindern angesprochenen Themenaspekten und zum anderen am Modell der Kausalattribution nach Weiner (1986).

Ein generelles Problem bei der Präsentation qualitativer Forschung besteht im Umfang des Datenmaterials, das nicht auf einzelne statistische Kennwerte reduziert werden kann. Um die Ergebnisse der Pilotstudie dennoch in aller Kürze zu berichten, wird eine thematisch gegliederte, zusammenfassende Darstellung der wich-

tigsten Gesprächsinhalte gewählt, die mit einigen wenigen Transkriptausschnitten illustriert wird (die Namen der Kinder sind geändert).

### 5.1 Spaß an Mathematik

Unterschiede zwischen Mädchen und Jungen zeigen sich bereits deutlich in der Eingangsphase der Gruppendiskussionen. Auf die Frage der Gesprächsleiterin "Macht Euch Mathe denn Spaß?" äußern sich die Mädchen verhalten, die Jungen hingegen durchweg begeistert:

Sandra: Ja, geht so ...

Ines: Mmh, ja, manchmal, manchmal so – manchmal so. ...

Johanna: Eigentlich macht's Spaß, bloß manchmal – manche Aufgaben finde ich

dann nicht mehr so gut.

Timo: Super, das ist super!

Alle: Ja!

Alex: Ja, sehr, sehr gut!

Timo: Das ist das Beste an der Schule!

Die Mädchen begründen ihre Position, dass Mathematik manchmal Spaß macht und manchmal nicht, im weiteren Gesprächsverlauf ausführlich mit der Art der Aufgaben. Dabei kristallisiert sich heraus, dass die Mischung von Rechenarten, vor allem wenn sie die Division enthält, als schwierig empfunden wird. Offensichtlich bevorzugen die Mädchen eher leichte Rechenaufgaben, die sie gut beherrschen:

Ines: Ich mag's gerne, wenn's nur plus ist, das geht nämlich ...

Sandra: Ja, ich auch, ich auch!

Ines: ... oder nur minus, dann ist das nämlich am einfachsten, sonst muss man

das nämlich immer anders stellen.

Carol: Ja, plus und minus, plus und minus, die beiden sind immer ganz toll,

finde ich, weil die sind ...

Sandra: Geteilt nicht und ... Carol: ... ganz einfach.

Johanna: Bloß wenn da 'ne Reihe plus und dann 'ne Reihe minus ist ...

Sandra: Aber, aber wir rechnen ja jetzt auch mit mal!

Johanna: ... dann verwechsle ich das immer. Wir rechnen ja jetzt auch mit mal.

Sandra: Oder geteilt.

Johanna: Oder alles drei und geteilt. Das mag ich nicht so gerne geteilt und mal.

Danach gefragt, welches Fach in der Schule am meisten Spaß mache, äußert die Mehrzahl der Mädchen und Jungen eine Präferenz für Sport, wobei die Jungen überlegen, ob Mathematik nicht noch besser sei. Interessant ist, dass auch zwei der neun Mädchen angeben, Mathematik sei ihr Lieblingsfach, wenn auch nach einigem Hin und Her: Nachdem ein Mädchen sagte: "Ich mag kein Rechnen", antwortete ein anderes sofort: "Ich auch nicht". Kurz darauf befragt, was ihr denn am besten gefalle, meinte sie zum Erstaunen der Gesprächsleiterin: "Rechnen". Diese Ambivalenz, die bei mehreren Mädchen zu beobachten ist, deutet darauf hin, dass sie in der Ausbildung ihrer Vorlieben und Abneigungen in Bezug auf den Fächerkanon der Grundschule noch weniger festgelegt sind als die befragten Jungen.

# 5.2 Mädchen und Jungen im selbst eingeschätzten Leistungsvergleich

Von den Mädchen sagen nur zwei, dass ihnen Rechnen leicht fällt. In der Jungengruppe behaupten alle, Rechnen sei "leicht, leicht, leicht". Auch auf komplexere Aufgaben angesprochen, bleiben die Jungen dabei, Mathe sei leicht. Dennoch ist ihnen Mathematik nicht langweilig, was man ja vermuten könnte, wenn alles so leicht ist.

Auf die Frage, welche Matheleistungen sie im Vergleich mit den anderen Kindern der Klasse haben, kam eine Mädchengruppe sehr schnell auf den Vergleich mit Jungen zu sprechen:

Ines: Ja, die meisten sind schneller als ich fertig.

Int.: Die sind schneller?

Sandra: Alex. Ines: Ja.

Es werden von den anderen noch weitere Jungennamen genannt, aber kein Mädchen.

Int.: Die Jungen sind so schnell?

Carol: Ja, die Jungen.

Johanna: Die müssen lahmer werden!

Als die Gesprächsleiterin später nachfragt, ob schneller denn auch besser heiße, nennen die Mädchen einige Jungen, die auch Fehler machen. Die andere Mädchengruppe einigt sich umgekehrt zunächst darauf, Mädchen und Jungen seien gleich gut in Mathematik, bringt aber später den Begriff Mädchenfach (z. B. Malen) ins Gespräch und benennt Rechnen als Jungenfach, in dem die Jungen besser seien. Auf diesen Widerspruch angesprochen, sagt ein Mädchen: "Ab und zu sind sie besser als wir" und ein anderes verallgemeinert auf "Die Jungen sind immer besser".

Beim selbst eingeschätzten Leistungsvergleich zeigt sich ein ähnliches Bild wie beim Thema Spaß an Mathematik: Die Jungen formulieren selbstbewusst einen festen Standpunkt, nämlich dass die Mädchen schlechter seien als sie. Die Mädchen stellen sich in ihrer Einschätzung wiederum ambivalent dar und vermuten zum Teil, sie seien schlechter als die Jungen. Nach Beurteilung der Lehrerin gibt es in ihrer Klasse *keine* Leistungsunterschiede in Mathematik zwischen Mädchen und Jungen.

# 5.3 Ursachenzuschreibungen

Erfolg oder Misserfolg, das heißt hier: nach eigener Einschätzung in Mathematik "gut" oder "schlecht" zu sein, führen die Kinder primär auf die Aufmerksamkeit im Unterricht zurück. Aufmerksamkeit heißt vor allem: richtig zuhören und nicht nebenbei reden, flüstern oder malen. Des Weiteren werden noch die Faktoren Üben, Abschreiben und Lust genannt. Abschreiben wird v. a. als Fehlerquelle erwähnt (man schreibt auch die Fehler mit ab); Lust wird als Motivationsfaktor eingeführt, der die Aufmerksamkeit beeinflusst. Wenn alle im Unterricht aufpassten, so sind sich die Kinder in allen drei Gruppen einig, "dann würden alle irgendwie gut sein".

Hinsichtlich des Vergleichs zwischen Mädchen und Jungen scheint kein Kind zu glauben, dass Mädchen oder Jungen qua Geschlecht besser oder schlechter in Mathematik sind. Auch Konzepte wie Fähigkeit oder Begabung werden von den Kindern nicht genannt – Konzepte, die ihnen vermutlich noch nicht zur Verfügung stehen (vgl. Nicholls 1978). Stattdessen werden folgende, internal variable Faktoren für die vermuteten Leistungsunterschiede hervorgehoben: Die Mädchen reden im Unterricht mehr; die Mädchen hören nicht richtig zu; die Mädchen schreiben zu viel ab; die Jungen üben besser. Die Lehrerin konnte allerdings die von den Kindern vorgetragene Auffassung, Mädchen wären im Mathematikunterricht weniger aufmerksam, würden mehr nebenbei reden und mehr abschreiben, ebenso wenig beobachten wie die vermuteten Leistungsunterschiede.

### 5.4 Der Einfluss der Peergruppe

Deutlich wird in unseren Gruppendiskussionen besonders ein Phänomen, nämlich der Einfluss der Peergruppe. Als Beispiel sei eine Gesprächspassage zitiert, in der die Gesprächsleiterin die Jungen fragt, ob denn die Mädchen auch irgendwo besser als die Jungen seien:

Alex: Im Rechnen nicht! ...

Markus: Religion ...

Marvin: Da bin ich aber auch gut! (empört)

Timo: In Religion sind die gut.

Markus: In Religion sind die besser als wir!

Alex: Ja.

Marvin: Nee, in Religion bin ich sehr gut!

Markus: In Religion sind die Mädchen aber besser! ...

Marvin: Aber ich bin fast so wie die Mädchen! (aufgebracht)

Nachdem Timo, Markus und Alex darauf bestehen, Mädchen seien in Religion besser, lenkt Marvin – obwohl noch aufgebracht – ein: "Aber ich bin fast so wie die Mädchen!" Auf solche Weisen bilden sich Gruppenmeinungen, die zur Ausbildung von Geschlechtsstereotypen sicherlich ebenso beitragen wie die in der Literatur diskutierten Erwartungshaltungen von Eltern und Lehrkräften. Marvin musste erfahren, dass man in seiner männlichen Bezugsgruppe Religion eher für eine Mädchensache hält.

#### 6 Fazit

Die Ergebnisse der vorgestellten Pilotstudie bestätigen im Wesentlichen die Ergebnisse der bereits vorliegenden quantitativen Forschung in diesem Feld. Wo also liegt der komplementäre Erkenntnisgewinn einer vergleichsweise aufwändigen und nicht mal annähernd repräsentativen qualitativen Studie? Dies soll abschließend in zwei Punkten diskutiert werden.

1. Qualitative Analysen erlauben u. a. die *Explikation von Ambivalenzen*, wie die oben bei den Mädchen beschriebenen. Dies ermöglicht einen Einblick in Prozesse

der selbstreflexiven und sozial-kommunikativen Herausbildung von Meinungen, Einstellungen und Werthaltungen. Standardisierte Untersuchungen verlangen demgegenüber eindeutige Antworten (auch auf Ratingskalen muss man sich "festlegen"). In solchen Untersuchungen bleibt zudem – trotz Erprobung – ungeklärt, ob alle Befragten die Formulierungen so verstehen, wie sie beabsichtigt waren (man denke z. B. an Probanden, deren Muttersprache nicht Deutsch ist). Dieses Problem sollte nicht unterschätzt werden; es spielt auch in der qualitativen Forschung eine wichtige Rolle, insbesondere in interkulturellen Forschungskonstellationen. Allerdings kann bei offenen Verfahren nachgefragt werden und es sind Präzisierungen und Beispiele möglich, die es im günstigen Fall ermöglichen, die Sicht der Befragten präziser zu rekonstruieren. Ein genereller Vorteil qualitativer Erhebungen liegt darüber hinaus darin, dass die untersuchten Probanden ihre Sichtweisen mit Bezug auf ihren je individuellen Erfahrungsbereich ausführlich explizieren können. Dadurch kommen ihre eigenen Relevanzsetzungen zur Sprache, die im Zuge der weiteren Analyse zu neuen (Sub-)Kategorien und/oder Hypothesen führen können.

2. Die vorgestellte Pilotstudie verweist auf einen Aspekt, der in der traditionellen Selbstkonzeptforschung bislang zu wenig beachtet worden ist, nämlich die Rolle der Peersozialisation bei der Entstehung unterschiedlicher Selbstkonzepte von Mädchen und Jungen und darüber hinausgehend bei allgemeinen Geschlechtsstereotypisierungen. Die unterschiedlichen Selbstkonzepte sind zwar statistisch zu einem relevanten Anteil auf differenzielle Eltern- und Lehrererwartungen zurückzuführen (vgl. Tiedemann 2000); die Rolle der Peersozialisation wurde diesbezüglich aber noch nicht systematisch untersucht. Dieser in theoretischer Hinsicht interessante Punkt könnte es sinnvoll erscheinen lassen, das Selbstkonzept einmal aus der Perspektive des Konzepts sozialer Repräsentationen zu untersuchen. "Social representations are shared and as such are properties of social groups and not isolated individuals" (Emler & Ohana 1993, S. 85). Hierfür scheint gerade die Methode der Gruppendiskussion ein geeignetes Forschungsinstrument zu sein, weil diese – im Idealfall – Einblicke in die Mikrogenese von Meinungen, Einstellungen und Werthaltungen und in die zugrundeliegenden sozialen Aushandlungsprozesse ermöglicht (vgl. Kölbl & Billmann-Mahecha 2005).

Aus all dem ergibt sich, dass die qualitative Forschung nicht nur der vorgängigen Hypothesengenerierung für die weitere quantitative Forschung dienen kann – eine Rolle, die ihr sowieso nie streitig gemacht wurde. Vielmehr kann sie, wie das Forschungsbeispiel ansatzweise zeigen sollte, eigenständige und zur quantitativen Forschung komplementäre, theoretisch relevante Beiträge leisten. In welcher Art von Design solche Beiträge geleistet werden und ob man die Zusammenführung dann Triangulation oder Mixed Methods nennt, scheint dabei zweitrangig. Wichtiger sind die Beendigung des "Paradigmenstreits" – ganz im Sinne von Wilhelm Kempf, dem dieser Aufsatz in freundschaftlicher Verbundenheit und mit kollegialem Respekt gewidmet ist – und die begründete Zurückweisung weder wissenschaftspraktisch noch wissenschaftstheoretisch haltbarer dichotomer Gegenüberstellung. Zu entwickeln wäre stattdessen eine beide Zugänge integrierende Forschungslogik, die dem Primat des Gegenstandes vor der Methode folgend der Vielfältigkeit und Komplexität des menschlichen Daseins gerecht wird.

# Subjektwissenschaftliche Theorie und Empirie

Morus Markard

# 1 Bezugspunkte in der Entwicklung des Methodenverständnisses der Kritischen Psychologie

Die hier in Rede stehende subjektwissenschaftliche Theorie und Empirie hat als kategorialen Bezug die von Klaus Holzkamp und anderen begründete Kritische Psychologie, die in erster Linie ein inhaltlicher Ansatz ist, dessen Ausarbeitung stetig mit methodologischen Analysen und Vorschlägen verbunden war. Wegen dieser Verschränkung von inhaltlicher und methodischer Entwicklung der Kritischen Psychologie zu einer Subjektwissenschaft ergibt sich die damit verbundene Bestimmung des Theorie-Empirie-Verhältnisses weder aus den Erfordernissen von einzelnen Methoden noch aus übergreifenden methodologischen Orientierungen (qualitativ/quantitativ), sondern aus der kategorialen (und daraus folgenden methodologischen) Eigenart psychologischer Forschung vom Standpunkt des Subjekts.

Zunächst soll nachgezeichnet werden, wie sich dieser Forschungsstandpunkt, der nicht metaphorisch gemeint, sondern wörtlich zu verstehen ist, mehr oder weniger sukzessive in der Reflexion von Desideraten vorfindlicher psychologischer Forschung gebildet hat.

Allgemeinster Bezugspunkt der Kritik ist dabei, was der Phänomenologe Carl-Friedrich Graumann den Primat der Methode vor dem Gegenstand bzw. den "Primat der experimentellen Methodik" (1979, S. 285f.) genannt hat, dass nämlich die Möglichkeit der im eigentlichen Sinne wissenschaftlichen Erforschbarkeit psychologischer Probleme davon abhängig gemacht wird, wieweit sie mit der experimentellen Anordnung durchführbar ist. Nun lässt sich zwar jede psychologische Problematik so entspezifizieren, dass sie in diese Anordnung passt; die Frage ist nur, um welchen Preis. Als Reflexion eben dieser Frage lässt sich die Entwicklung der kritisch-psychologischen Methodenkritik verstehen, an der ich zwei Aspekte herausheben will.

# 1.1 "Organismische Anthropologie" der experimentellen Forschungspraxis

Erstens: Klaus Holzkamp war schon etablierter Experimentalpsychologe, als er die, wie er es nannte, "verborgenen anthropologischen Voraussetzungen" der experi-

166 Morus Markard

mentellen Vorgehensweise explizierte. Im Verhältnis von Versuchsperson (Vp) und Versuchsleiter(in) (Vl) werde die Möglichkeit einer dialogischen Beziehung von vorneherein unterlaufen (1972b, S. 47). Die experimentelle Forschung gehe "von der Idee einer Art 'Norm-Vp'" aus, die ein bloß "gedachtes" Individuum sei, das sich an undurchschaute "Umweltbedingungen" anzupassen habe (S. 58), womit Menschen dazu gebracht werden sollten, "sich wie 'Organismen' zu verhalten" (S. 61). Insofern liege dieser Forschungspraxis eine organismische Anthropologie zugrunde. Diese Kritik zielt vor allen darauf, dass das Verhältnis von objektiver Bestimmtheit und subjektiver Bestimmung, das wesentlich für menschliche Existenz ist, um die subjektive Bestimmung verkürzt ist. Allerdings ist in der experimentell-psychologischen Anordnung mit der Formulierung von Bedingungs-Ereignis-Relationen nur das viel allgemeinere Wirkungs- oder Bedingtheitsdenken in der Psychologie und in den Sozialwissenschaften auf den praktisch elaboriertesten Begriff gebracht, das insoweit von der Experimentalkritik mit betroffen ist. <sup>1</sup>

# 1.2 Operationalisierungsprobleme bei Experimenten und Umfragen und das "Zahnpastatuben-Problem"

Der zweite Aspekt thematisiert die Diskrepanz zwischen dem "Kanon statistischer Prüfmethodik einerseits und der ... Ungenauigkeit bei der begrifflichen Fassung dessen, was da eigentlich untersucht werden soll" (Holzkamp 1981, S. 276f., Hervorhebung im Original)<sup>2</sup>. Man könnte auch sagen, dass Operationalisierungen weiter reichende theoretische Klärungen ersetzen – wie übrigens auch in der Meinungsforschung, wie sich an der Antiamerikanismus-Messung in Band 8 der von Heitmeyer herausgegeben Reihe Deutsche Zustände veranschaulichen lässt

Für eine mögliche Nutzung experimenteller Anordnungen in subjektwissenschaftlicher Forschung (Markard 2009, S. 53ff.) ist der Ausgangspunkt der von Holzkamp (1970, S. 39) aufgegriffene Gedanke Habermas' (1965, S. 1147), dass psychologische Gesetzmäßigkeiten (bzw. eher "spontane Tendenzen", wie etwa Verantwortungsdiffusion in Hilfesituationen, vgl. Markard 2009, 52f.) zwar nicht außer Geltung, aber außer Anwendung gesetzt werden können, weil und soweit die Menschen sich diese Tendenzen bewusst machen und sich zu ihnen verhalten können. Entsprechend ginge es darum, Situationen herzustellen, in denen sich diese Tendenzen realisieren. "Es stünden dann weniger die nackten Resultate des Experimentes zur Diskussion; vielmehr ginge es um die systematische Analyse der im Experiment gemachten Erfahrungen, der an den Tag gelegten Handlungen. Was bei experimentellen Untersuchungen als postexperimentelle Befragungen fungiert, wäre unserer Konzeption nach der eigentliche Beginn der Forschung, und nicht bloß eine Art Anreicherung der experimentell gewonnenen Daten" (ebd., S. 55). Das nicht auflösbare Problem dabei ist, dass partiell die Mitforschungs-Beziehung unterlaufen werden muss.

Dabei ist auch zu bedenken, dass generell mit den üblichen empirischen Methoden – qualitative eingeschlossen (!) – nicht über die wissenschaftlich Relevanz von Begriffen, in denen zu prüfende Theorien formuliert sind, entschieden werden kann: So stellt das Reiz-Reaktions-Lernen eine Art und Weise dar, sich der Umwelt anzupassen, die wir Menschen – als Möglichkeit – mit vielen Arten teilen. Bedeutungsvermitteltes Lernen dagegen ist spezifisch für den Menschen: Wenn das Aufleuchten eines Bremslichts für Menschen ein bloßer Reiz wäre, müssten Fahrschüler(innen) seine Funktion über eine Reihe von Auffahrunfällen lernen; erfreulicherweise kann man Menschen aber die Bedeutung des Bremslichts verbal vermitteln. Beide Lernarten – Reiz-Reaktions- und Bedeutungslernen – sind empirisch zu realisieren. Die Klärung ihrer Relevanz allerdings kann auf diese Weise aber eben nicht erfolgen. Der Vorschlag der Kritischen Psychologie, historisch-empirisch an das Begriffsproblem der Psychologie heranzugehen, kann hier nicht referiert werden (vgl. zu einer Übersicht Markard 2009, S. 92ff.).

(Knappertsbusch & Kelle 2010), Nachdem dort zunächst die Unschärfe des Konzeptes Antiamerikanismus ventiliert wird und man sich von der "Lösung der begrifflichen ... Problematik weit entfernt" sieht (ebd., S. 159), wird das Konzept gleichwohl in vier Items operationalisiert, die (als Teil einer Itembatterie in ca. halbstündigen Telefoninterviews) abgefragt und dann interpretiert werden. Die Befragten sind ebenso reine Datenlieferanten wie die Vpn im Experiment. Das ist - wohlgemerkt - keine Kritik speziell an dieser Untersuchung, sondern eine kritische Explikation des Standards auf Einstellungsmessung abzielender Untersuchungen (vgl. dazu auch Markard 1984). Der in diesem Kontext geäußerten "gualitativ" inspirierten Kritik, hier falle man hinter den Alltag zurück, stehe ich deswegen skeptisch gegenüber, weil mir der bürgerliche Alltag nicht gerade von herrschaftsfreiem Diskurs sensu Habermas bestimmt zu sein scheint. Die – kontrafaktische – Idealsetzung des herrschaftsfreien Diskurses zielt auf Kommunikationsverzerrungen im Alltag und in reduktiver Methodik – auch wenn Adorno meinte, der Abfragestil der Meinungsforschung passe zu den heruntergekommenen Befragten: "Dort, wo die Menschen unter dem Druck der Verhältnisse in der Tat auf die 'Reaktionsweise von Lurchen' [hier verweist er auf die *Dialektik der Aufklärung*, M.M.] heruntergebracht werden, wie als Zwangskonsumenten von Massenmedien und anderen reglementierten Freuden, passt die Meinungsforschung, über welche sich der ausgelaugte Humanismus entrüstet, besser auf sie als etwa eine 'verstehende' Soziologie: denn das Substrat des Verstehens, das in sich einstimmige und sinnhafte menschliche Verhalten, ist in den Subjekten selbst schon durch bloßes Reagieren ersetzt" (1957, S. 202f.)<sup>3</sup>. Der Unterschied zu Holzkamps oben geschilderter Argumentation ist der, dass Holzkamp (der diese Argumentation Adornos nicht kannte) die organismische Anthropologie als Denkform, nicht aber als Realität ansieht.

Während im Experiment die für die Interpretation erforderlichen Bezugsgrößen im Setting hergestellt und somit praktisch definiert sind, trifft dies bei Meinungserhebungen nicht zu, so dass für Interpretationen auch weniger gesichertes Bezugsmaterial vorhanden ist. Jedenfalls können im interpretativen Paradigma Korrelationen "*nicht* als statistisch-induktive *Erklärungen* hingenommen werden, sondern lediglich als *Beschreibungen* eines *erklärungsbedürftigen Sachverhalts*" angesehen werden (Kempf 1992a, S. 106, Hervorhebung im Original), wobei Signifikanz nur darauf verweise, dass beobachtete Regelmäßigkeiten einer weiteren Beachtung würdig seien. Die statistische Verteilung von Sinngehalten sieht Kempf als

Die Befragten werden faktisch gezwungen, knapp zu regieren – wobei einzuräumen ist, dass dies auch Anforderungen im Alltag entsprechen kann, etwa bei Wahlen, bei denen potenziell differenzierte Sinnzusammenhänge auf ein Kreuzchen zusammenschnurren. Die Frage ist nur, ob Wissenschaft das reproduzieren soll. Zur Veranschaulichung dessen, wie sehr die Methodenwahl die Ergebnisse beeinflusst, ist m. E. nach wie vor die klassische Untersuchung LaPieres einschlägig, der Mitte der dreißiger Jahre des vorigen Jahrhunderts mit einem chinesischen Ehepaar quer durch die USA reiste und aus Fragebogenuntersuchungen heraus Vorurteile und Diskriminierung erwartete. Aber: Mit einer einzigen Ausnahme wurden sie bedient, auch in den 31 Fällen, in denen die Chinesen ohne LaPiere eintraten – in einem Drittel der Häuser sogar besonders zuvorkommend. Sechs Monate später wurde brieflich die Frage gestellt, ob Chinesen akzeptiert würden (auch an vorher nicht besuchte Einrichtungen): Es gab einen 50%igen Rücklauf und 90% Ablehnungen (vgl. Markard 1984, S. 105ff.).

168 Morus Markard

Voraussetzung dafür, die Verbreitung etwa antisemitischer Vorurteile zu erheben; deren *Zustandekommen* betrachtet er allerdings als nur qualitativ untersuchbar.

Die methodisch induzierte Enthistorisierung und Entkonkretisierung subjektiven Sinns in Experiment und Umfrage ist nicht nachträglich wieder hineinzuinterpretieren. Das ist das Zahnpastatuben-Problem: Was man methodisch rausgedrückt hat, kriegt man auch theoretisch nicht wieder rein.

### 1.3 Das subjektwissenschaftliche Minimalprogramm

Ex negativo ergibt sich aus dem Gesagten folgendes methodische Minimalprogramm: *Erstens:* Objektive Bestimmtheit *und* subjektive Bestimmung menschlicher Existenz müssen berücksichtigt werden; eine Überlegung, die in Richtung Handlungsforschung, die Einheit von Erkennen und Verändern, geht. Dies passiert zwar auch im Experiment, in dem Erkenntnis sich ja über die praktische Veränderung der Variablen ergibt. Handlungsforschung i. e. S. bedeutet aber, dass die Betroffenen aktiv in ihre Lebensumstände eingreifen. *Zweitens:* Beides schließt eine gewisse Symmetrie zwischen den Beteiligten ein, die zum Forscher-Mitforscher-Verhältnis zu entwickeln ist. *Drittens:* Die inhaltliche Bedeutungshaftigkeit menschlicher Welt- und Selbstbegegnung ist zu berücksichtigen, wobei das, *was* da als bedeutsam erfahren wird, z. T. interdisziplinär, gesellschaftstheoretisch zu erfassen ist. Es wird also keine weitgehende Freiheit von Vorannahmen postuliert, sondern diese werden – im Benehmen mit den Mitforscher(inne)n – formuliert und in den Forschungsprozess eingebracht.

# 2 Die Kategorie Handlungsfähigkeit und die Gesellschaftlichkeit menschlicher Existenz

Diesem Programm entspricht, dass Handlungsfähigkeit die zentrale Kategorie der Kritischen Psychologie ist, und zwar Handlungsfähigkeit in den Widersprüchen unserer Gesellschaft, Gesellschaftliche Lebensbedingungen sind aber nicht als unmittelbare Aktivitätsdeterminanten psychologisch relevant, sondern als Bedeutungen, zu denen sich die Individuen verhalten können und müssen. Wegen der Historizität menschlichen Erlebens und Verhaltens ist die gesellschaftstheoretische Analyse gesellschaftlicher Lebensbedingungen psychologisch unverzichtbar, "ihre konkrete psychologische Bedeutung hat sich aber vom Standpunkt des Subjekts aus zu erweisen" (Markard 2009, S. 151, Hervorhebung im Original). Die Individuen haben unter Herrschaftsverhältnissen grundsätzlich die doppelte Möglichkeit, sich mit bestehenden Möglichkeiten restriktiv zu arrangieren oder diese Möglichkeiten (zusammen mit anderen) auszuweiten. Das damit auch bestehende doppelte Problem, im restriktiven Arrangement auf Lebensmöglichkeiten zu verzichten bzw. mit der Erweiterung von Lebensmöglichkeiten Konflikte (in Richtung auf verallgemeinerte Handlungsfähigkeit) zu riskieren, markiert die Grundkonstellation aktual-empirischer kritisch-psychologischer Forschung zwischen objektiver Bestimmtheit (durch gesellschaftliche Verhältnisse) und subjektiver Bestimmung

(als Möglichkeit des Eingriffs in diese Verhältnisse). In der Art und Weise und soweit ein Individuum Bedingungen und deren Bedeutungen *für sich* als Handlungsmöglichkeiten bzw. -behinderungen wahrnimmt und akzentuiert, macht es sie für sich zu seinen Prämissen. Prämissen sind folglich nicht bloß Aspekte von Bedeutungskonstellationen, sondern individuell und in subjektiven Lebens-Interessen begründet aus diesen "herausgegliedert".

In diesem Sinne ist die Analyse von Bedingungen und Bedeutungen bei der Klärung psychologischer Fragen und Probleme unverzichtbar, will man die Handlungsprämissen und Handlungsgründe der Einzelnen verstehen. Diese Bedingungen können aber nicht – soziologistisch – einfach auf das individuelle Handeln, Denken und Empfinden "herunterkonkretisiert" werden. Das heißt: Der psychologische Erkenntnisweg ist nicht der einer zunehmenden Konkretisierung allgemeiner gesellschaftlicher und dann institutioneller Bedingungen auf das Problem hin, sondern umgekehrt der Weg von ungelösten Aspekten des Problems hin zu Bedingungen, die für dessen Analyse und Lösung von Bedeutung sein können. Die kritisch-psychologische Bedingungen aufzuschließen sind, ist eine Frage des jeweiligen Forschungsprozesses und nicht unabhängig von ihm zu beantworten.

Ein so angelegter Forschungsprozess muss die Trennung zwischen Theorie und Empirie so aufheben, dass in der Theoriebildung der Bezug zur Realität schon angelegt ist, dass 'Nichtprofessionelle' als Mitforschende verstanden werden, mit denen gemeinsam ihre Welterfahrungen zum Gegenstand zu machen und zu reflektieren sind.

# 3 Begründungsdiskurs als Wissenschaftssprache der Psychologie

Damit ist folgende Charakterisierung der Theoriesprache bzw. von Theorien verbunden: Theorien sollen zur Selbstverständigung der Beteiligten dienen; eine Intention, die als subjektwissenschaftlich zu charakterisieren ist: In diesem Begründungsdiskurs geht es um die Welt, wie die Menschen sie erfahren, und um die konkrete Klärung des erwähnten Verhältnisses von objektiver Bestimmtheit und subjektiver Bestimmung. Anders formuliert: Statt um Bedingungs-Ereignis-Relationen (Ursache-Wirkungs-Zusammenhänge) geht es um Prämissen-Gründe-Zusammenhänge. Die Vermittlung gesellschaftlicher Bedeutungen mit individuellen Handlungen vollzieht sich im Begründungsdiskurs. Die Spezifik von Handlungsbegründungen gegenüber der unvermittelten Wirkung von Bedingungen (Bedingtheitsdiskurs) liegt darin, dass "Begründungen nur vom Standpunkt des Subjekts aus möglich sind: Gründe sind immer 'erster Person', d. h. 'je meine"Gründe" (Holzkamp 1991, S. 6, Hervorhebung im Original). Beim Begründungsdiskurs haben wir es weder mit einer "Methode" noch mit "einer 'Theorie' zu tun, ... sondern ... mit einer bestimmten *Diskursform* intersubjektiven Umgangs ... Entsprechend liegt die einzige Möglichkeit, den Begründungsdiskurs in seiner Besonderheit zu qualifizieren, darin, ihn vom Bedingtheitsdiskurs, dessen Nexus nicht als Bedeu170 Morus Markard

tungs- / Begründungszusammenhang, sondern als 'Ursache-Wirkungs-Zusammenhang' spezifiziert ist, abzuheber" (Holzkamp 1996, S. 64, Hervorhebung im Original). Kempf hat diesen Grundgedanken auf das "intentionalistische Erklärungsschema" und die dem "Handlungsbegriff inhärente Vernünftigkeit der Akteure" bezogen – als "methodisches Prinzip, unter welchem das beobachtbare Verhalten als sinnrationales Handeln darstellbar und damit argumentationszugänglich wird" (1994b, S. 55f., Hervorhebung im Original).

Sofern man das Subjektivitäts-Objektivitäts-Verhältnis als Verhältnis gegenseitiger Ausschließung ansieht, sind methodische Vorkehrungen zu treffen, subjektive Eingriffe zu eliminieren oder zu kontrollieren. Aus subjektwissenschaftlicher Sicht allerdings handelt es sich bei derartigen methodischen Vorkehrungen zur Sicherung wissenschaftlicher Objektivität um die Herstellung eines künstlichen zwischenmenschlichen Mangelzustandes, der soziale Selbstverständigung desavouiert – oder, wie es Adorno (1972, S. 69) formulierte, "Erfahrung annulliert".

## 4 (Verborgener) Begründungsdiskurs in nomothetisch sich verstehender Psychologie

Eine methodologische Pointe der Argumentation mit dem Begründungsdiskurs liegt nun darin, dass sich dieser auch in nomothetisch formulierten und gemeinten Theorien findet, wie Holzkamp am Beispiel von sozialpsychologischen Theorien (1986) und Lerntheorien (1993) gezeigt hat. Wenn die Beziehung zwischen dem Wenn- und dem Dann-Teil einer Hypothese oder Theorie sinnvermittelt ist, handelt es sich – aller nomothetischen Rhetorik zum Trotz – nicht um eine Bedingungs-Ereignis-Konstellation, sondern einen Prämissen-Gründe-Zusammenhang bzw. ein verborgenes Begründungsmuster. Dies lässt sich über die Frage klären, ob es möglich ist, zwischen den Wenn- und den Dann-Teil einer psychologischen Aussage ein "subjektiv vernünftigerweise" oder "subjektiv funktionalerweise" zu schieben. Vernünftigerweise steht dabei nicht für ein externes Rationalitätskriterium, sondern allein für subjektive Begründetheit und Funktionalität. Inwieweit diese subjektive Begründetheit auch externen Rationalitätskriterien entspricht oder nicht (etwa bei spieltheoretischen Fragestellungen), ist für den Begründungsdiskurs eine nachgeordnete Frage: Dass menschliches Handeln in Prämissen begründet ist, hängt nicht davon ab, ob es auch nach subjekt-externen Kriterien rational ist. Wenn es nun stimmt, dass in nomothetisch gemeinten Theorien Prämissen-Gründe-Zusammenhänge enthalten sind, kann der Begründungsdiskurs nicht in eine hermeneutische Exklave abgeschoben werden, sondern er zieht sich durch die gesamte Psychologie.

In einer Diskussion *Zur Problematik des Empiriebezugs psychologischer Theorien* hat Wilhelm Kempf das holzkampsche Kriterium, vorfindliche Theorien durch den Einschub "vernünftigerweise" auf verborgene Begründungsmuster hin zu analysieren, als "zwar sehr wesentlich" bezeichnet, "aber es begründet, glaube ich, nur einen Verdacht und noch nicht definitiv den Nachweis, dass hier wirklich eine Be-

gründungstheorie vorliegen muss. Ein evolutionsbiologisch selegiertes, funktionales Verhalten erscheint ja auch vernünftig, so dass zumindest noch eine Trennung zwischen 'funktionalerweise' und 'vernünftigerweise' vorzunehmen wäre" (1994c, S. 19)<sup>4</sup>. Diese terminologische Unterscheidung beugt rationalistischen Fehldeutungen subjektiver Vernünftigkeit sicherlich vor. Es scheint mir bei Kempfs Einwand aber mehr um unbewusste Prozesse zu gehen bzw. darum, dass subjektive Funktionalität etwa evolutionsbiologisch gedeutet werden kann. Ein Prämissen-Gründe-Zusammenhang muss aber keineswegs bewusst sein. Die Rede vom Unbewussten als psychischer Dynamik macht vielmehr erst im Begründungsdiskurs Sinn. Im Bedingtheitsdiskurs ist dagegen weder die Rede vom Bewussten noch vom Unbewussten sinnvoll: Lackmus-Papier färbt sich zwar gewiss nicht bewusst blau oder rot, aber eben auch nicht unbewusst, sondern schlicht bedingt. Der Umstand, dass Prämissen-Gründe-Zusammenhänge unbewusst sein können, macht es überhaupt erst erforderlich, sie in (sozialer) Selbstverständigung aufzuschlüsseln. Allerdings werden die Grenzen des Begründungsdiskurses deutlich, wenn wir uns physiologienahen Bereichen nähern. So ist der Zusammenhang von Alter und Gedächtnisleistung sicher nicht subjektiv funktional; er kann für mich aber zur Prämisse werden, bestimmte Strategien gegen Vergesslichkeit (wie die berühmten Taschentuchknoten [aus der Vor-"Tempo"-Zeit, wenn ich mich recht entsinne]) zu entwickeln. Bei evolutionsbiologischen Bezügen wäre zunächst zu fragen, inwieweit diese biologistisch sind, also gesellschaftliche Funktionalitäten bzw. Anrufungen auf die Natur projizieren. Sofern dies ausgeschlossen werden kann, befinden wir uns tatsächlich an der Grenze des Begründungsdiskurses bzw. da, wo die Frage auftaucht, ob und und wenn ja, wie derartige Funktionalitäten für mich zu – unbewussten – Handlungsprämissen werden.

## 5 Datenbezüge und Verallgemeinerung in subjektwissenschaftlicher Forschung

Methodologisch zentral ist nun der Umstand, dass es sich bei einem Prämissen-Gründe-Zusammenhang um eine i. w. S. definitorische Bestimmung 'vernünftigen' Verhaltens handelt. In diesem Sinne sind – rein formal gesehen – Prämissen-Gründe-Zusammenhänge implikativ und einer empirischen Prüfung weder bedürftig noch fähig, so dass entsprechende Prüfungen pseudo-empirisch sind. Bei Prämissen-Gründe-Zusammenhängen haben empirische Daten nicht die Funktion, Zusammenhangsaussagen (Hypothesen und Theorien) zu prüfen, sie können sie nur veranschaulichen bzw. konkretisieren. Grundsätzlich aber markiert es einen bemerkenswerten Widerspruch der experimentell-statistisch orientierten Psychologie, dass sie *theoretisch* durchaus subjektive Handlungszusammenhänge thematisiert, diese dann aber *methodisch* unterschreitet. Möglich ist das, wie oben schon angedeutet, weil es praktisch keinen Aspekt menschlicher Handlungen gibt, der *nicht* variablenpsychologischer Bedingungsanalyse zugänglich gemacht werden könnte.

<sup>4 2003</sup> hat Kempf in Form einer Tabelle eine Reihe von Variablenzusammenhängen aufgestellt (2003a, S. 168f., Tab. 2.5.6) und danach unterteilt, ob sie für Sinngehalte "anfällig" sind.

172 Morus Markard

Es ist praktisch immer möglich, Prämissen-Gründe-Zusammenhänge auf Bedingungs-Ereignis-Relationen zu reduzieren. Das Problem ist nur, dass man von der Handlung zwar zum Verhalten kommen kann, aber nicht mehr von da aus zurück: das erwähnte Zahnpastatuben-Problem. Worüber im Kontext *experimentell*-statistischer Anordnungen dann sinnvoll "spekuliert" werden kann, sind m. E. *nicht* die statistischen Verteilungen, sondern die davon völlig unabhängigen, theoretisch gesetzten Prämissen-Gründe-Zusammenhänge. Und zwar deswegen, weil in Verteilungen die subjektive Sinneinheit des einzelnen Falles zerstört wird, während Prämissen-Gründe-Zusammenhänge auf einzelfallbezogene Strukturaussagen zielen.

Da subjektwissenschaftliche Aussagen die Form von Prämissen-Gründe-Zusammenhängen besitzen, können damit weder Bestimmungen zur Häufigkeit der in ihnen behandelten Phänomene vorgenommen werden, noch sind sie durch beliebig viele Fälle zu beweisen oder zu widerlegen. Sie sind aber relevant für die Klärung der Frage, ob dieselben Prämissen-Gründe-Zusammenhänge auch bei anderen Menschen vorliegen oder ob Prämissen spezifiziert werden müssen.

Die Charakterisierung subjektwissenschaftlicher Geltung und Verallgemeinerung bezieht sich auf praktische Lebensvollzüge der Individuen in historisch-konkreten Konstellationen, auf subjektive Möglichkeitsräume oder Handlungsmöglichkeiten – nicht auf Merkmale. Holzkamp (1983, S. 545) hat hierfür den Begriff Möglichkeitsverallgemeinerung vorgeschlagen, in der die subjektive Befindlichkeit bzw. (begrenzte) Handlungsmöglichkeit als "*Verhältnis zwischen allgemeinen geselschaftlichen Handlungsmöglichkeiten* und *meiner besonderen Weise ihrer Realisierung*" oder Einschränkung (S. 548, Hervorhebung im Original) zu begreifen ist.

### 6 Psychologie vom Standpunkt des Subjekts

"Psychologie vom Standpunkt des Subjekts" ist also tatsächlich nicht metaphorisch gemeint, sondern in dem Sinne wörtlich zu nehmen, dass die Menschen *nicht be*forscht werden, sondern auf der Forschungsseite stehen sollen. Gegenstand der Forschung sind nicht (andere) Menschen, sondern die Welt, wie sie von ihnen erfahren wird (anstatt dass danach gefragt wird, wie die Welt auf die Individuen wirkt).

Verstehen und Interpretieren oder die Gewinnung der Innensicht des Subjekts allein bedeuten allerdings keineswegs, *kein* Herrschafts- und Kontrollinteresse zu bedienen; und darin, inwieweit über die Köpfe der Betroffenen hinweg interpretiert wird, unterscheiden sich quantitative und qualitative Methoden nicht *per se*; sie unterscheiden sich nur in der Art der Daten, die in die Interpretation Eingang finden können. Auch das Verstehenwollen des Subjekts kann Ausdruck von Kontrollinteresse sein.

Die Perspektive einer Psychologie vom Standpunkt des Subjekts ist insofern ein Regulativ dagegen, als hier die Interessen der Beteiligten offen gelegt werden sollen – und müssen, wenn eine den Modus der intersubjektiven Verständigung einschlie-

ßende Forschungsbeziehung gelingen soll. Entsprechend darf in einem derartigen Forschungsprozess das intersubjektive Beziehungsniveau nicht unterschritten werden, es muss allerdings selbst Gegenstand von Reflexion in Richtung auf wissenschaftliche Geltung und Verallgemeinerbarkeit der erzielten Resultate sein. Holzkamp hat diesen Aspekt subjektwissenschaftlicher Forschung als Metasubjektivität bezeichnet.

### 6.1 Konzept der Entwicklungsfigur

Was dies bedeutet, soll abschließend am Konzept der Entwicklungsfigur erläutert werden, die – idealtypisch – einen subjektwissenschaftlichen Forschungsprozess als Folge von vier Schritten ("Instanzen") fasst (vgl. auch Markard 2010):

- 1. *Ausgang* von einem lebenspraktischen Problem (von Betroffenen), in welchem sich unbegriffene Machtverhältnisse niederschlagen.
- 2. Analyse der Problemlage: Ziel ist, die vorfindliche und strittige Vielfalt von Interpretationen der Problemlage, damit auch der kategorialen Grundlagen und subjektiven Theorien, in einer Forschungsgruppe einerseits zum Zuge kommen zu lassen, sie aber andererseits auch so zu reduzieren, dass alternative Handlungsmöglichkeiten herausgearbeitet werden können. In der zweiten Instanz der Entwicklungsfigur geht es also darum, diejenigen Prämissen-Gründe-Zusammenhänge herauszupräparieren, in denen die restriktive Funktionalität der erfolglosen Versuche der Bewältigung der thematisierten Probleme durch die Betroffenen verständlich wird. Da in diesen Debatten alle Daten zugelassen sind, müssen sie auf ihre Funktion und Modalität hin analysiert werden (vgl. ebd., S. 175).
- 3. Herausarbeitung von Handlungsvorschlägen: Ziel ist die Akzentuierung von Handlungsprämissen, unter denen die Betroffenen andere als die bisherigen Handlungsmöglichkeiten sehen bzw. Handlungsblockierungen auflösen können. In dieser Instanz soll also überlegt werden, wie das Ausprobieren von Neuem, von Alternativen praktisch vonstattengehen könnte.
- 4. *Rückmeldung und Auswertung der praktischen Erfahrungen:* Die Frage ist hier, ob die Analysen und Handlungsvorschläge praktisch relevant waren, ob sie modifiziert werden mussten, oder ob die Beteiligten in ihren Änderungsbemühungen scheiterten. Für diesen letzten Fall sind im Konzept der *Stagnationsfigur* bezüglich aller Instanzen Fragen zu den möglichen Gründen eines Scheiterns versammelt.

Gerade an der vierten Instanz lässt sich der idealtypische Charakter der Instanzenabfolge verdeutlichen: Im wirklichen Lebens- und Forschungsprozess wechseln sich Analysen, Erprobungen und Rückmeldungen immer wieder ab. Wichtiger noch: Weil in den stream of behavior eingebunden, lässt sich eine Entwicklungsfigur nicht wirklich abschließen (vgl. dazu auch Kempfs Überlegungen zu linearen und kreisförmigen Handlungsabläufen (1994b, S. 56ff.). Allemal aber gilt in diesem

174 Morus Markard

Zusammenhang: "Die Richtigkeit einer Handlungserklärung erweist sich dann am Erfolg der Konfliktlösungsbemühungen" (ebd., S. 60).<sup>5</sup>

### 6.2 Forschungsprobleme

Zwei kürzlich erschienene Dissertationen über den Umgang von Sozialarbeiter(inne)n mit neoliberalen Regulationen (Eichinger 2009) und über die Sicht von einsitzenden Intensivtätern auf ihre Situation (Huck 2009) zeigen exemplarisch, dass eine derartige Forschung kaum bruchlos realisierbar ist. Gerade die Arbeit von Huck fand unter besonders restriktiven Bedingungen statt. Als Hauptmangel seiner Untersuchung sieht er natürlich, dass es ihm nicht gelungen sei, eine kontinuierliche und intensive Mitforscher(innen)-Beziehung zu den Befragten aufzunehmen. Im Ergebnis habe er in der Auswertung der Interviewtranskripte teilweise rekonstruktiv vorgehen müssen, obwohl vielleicht die Möglichkeit bestanden hätte, Begründungszusammenhänge zumindest mit einigen Befragten gemeinsam zu formulieren (was aber mehrere Treffen erfordert hätte: dies war organisatorisch schwierig angesichts der Inhaftierung, außerdem nicht prioritär seitens des Auftraggebers). Trotz dieser Einschränkungen sieht Huck in seiner Untersuchung einen Erkenntnisgewinn: Teilweise auf die verbalen Einlassungen der Betroffenen bezogen, teilweise direkt aus diesen übernommen, ließen sich hypothetische Prämissen-Gründe-Zusammenhänge formulieren, die mit bestimmten Lebensumständen der Betroffenen zu vermitteln seien. Ziel der Auswertung war es, im Zuge der angesprochenen Möglichkeitsverallgemeinerung Konstellationen gesellschaftlicher Handlungsmöglichkeiten und Behinderungen zu identifizieren, die problematische Handlungen der Jugendlichen subjektiv notwendig/funktional machten. Der Autor schlägt dazu drei Aussagetypen (2009, S. 111ff.) vor: (1) "Faktische Aussagen", die, um psychologisch relevant zu werden, weiterer begründungstheoretischer Klärung bedürften; etwa der Umstand, dass in mehreren Fällen allein erziehende Mütter ihr Kind langfristig in ein Heim abgaben, in dem die Betroffenen durchgehend schlechte Erfahrungen machten. (2) "Rekonstruierte Begründungszusammenhänge", die angesichts realer Gegebenheiten mögliche und folgerichtige Prämissen-Gründe-Zusammenhänge formulierten: Das Wissen darum, dass nur wenige Hauptschüler(innen) Ausbildungsplätze finden, kann zur Prämisse einer Schulverweigerung werden; offen bleibt, auf welchen Fall dies passt. (3) "Subjektive Theorien", die von den Befragten selbst in die Diskussion eingebracht worden seien. Dies bedeute allerdings weder, dass diese unanzweifelbar seien, noch, dass man die Wertmaßstäbe teilen müsse, mit denen sie verbunden seien: So verstünden Jugendliche zwar die Norm- und Sanktionssysteme des Jugendgefängnisses, seien aber durch informelle soziale Regeln unter den Gefangenen zu Überschreitungen der formellen Regeln gezwungen, obwohl sie um die Konseguenzen wüssten: So sei es z. B. notwendig, sich in gewalttätigen Konflikten "wehrhaft" zu zei-

Da das, was sich hier lebenspraktisch bewährt (oder eben nicht), die Realisierung von (Theorien in Form von) Prämissen-Gründe-Zusammenhängen (Begründungsmustern) ist, haben die empirischen Daten, wie in Kap. 5. skizziert, keine Prüf-, sondern Veranschaulichungs- bzw. Konkretisierungsfunktion (vol. zur Entwicklung unserer diesbezüglichen Überlegungen Markard 2000, S. 239ff.).

gen, um nicht ständig angreifbar zu sein. Exemplarisch seien hier einige verschiedene Motive zu Straftaten wie Lebensstil, Langeweile, Ansehen erwähnt, aber auch die teilweise noch wirkenden Vorstellungen der Jugendlichen, wie sie sich der Strafverfolgung entziehen könnten, wie sie die ihnen drohenden Gefahren herunterspielten und dadurch ihre Ängste "abwehrten", etc.

Wie diese Arbeit zeigt, gilt nicht das Prinzip des "Alles-oder-nichts", bei dem nämlich nur "nichts" herauskäme und kritisch-psychologische Forschung zu einer normativen Nullnummer würde. Die geschilderten subjektwissenschaftlichen Forschungsregulative haben nicht den Sinn, Forschung zu verhindern, sondern zu optimieren bzw. Reflexionen darüber zu ermöglichen, welcher Erkenntnismöglichkeiten man sich begibt, soweit die geschilderten methodologischen Vorstellungen oder Regulative nicht realisiert werden können bzw. wenn man sie einfach nicht realisieren will – was ja jedermanns und jeder Frau Recht ist.

### Das argumentative Rückkoppelungsgespräch

Wassilios Baros & Klaus-Dieter Reetz

"Die Erkenntnis der Wirklichkeit ist ja nicht Selbstzweck, sondern hat ein einleuchtendes praktisches Ziel. Sie dient der Orientierung des Menschen in seiner Umwelt und der Absicherung des Erfolgs seines Handelns. Deshalb besteht ein praktisches Interesse daran, zwischen dem, was wirklich ist, und dem, was unwirklich ist, zu unterscheiden. Aus denselben Gründen besteht auch ein praktisches Interesse daran, zwischen dem zu unterscheiden, was wahr, und dem, was unwahr ist."

Kempf 2009b, S. 59

### 1 Einleitung

Das Verfahren des argumentativen Rückkoppelungsgesprächs versteht sich als forschungsmethodische Alternative zu konsensorientierten Ansätzen der kommunikativen Validierung. Interpret und Interpretierte nehmen zu den im Laufe des Forschungsverlaufs entwickelten Interpretationen argumentativ Stellung, kommunizieren ihre eigenen Wahrnehmungen und subjektiven Wirklichkeitskonstrukte, handeln intersubjektiv geteilte Wirklichkeiten aus und erarbeiten im Rahmen eines emanzipatorischen Diskurses Kriterien für die Überprüfung der Angemessenheit ihrer Handlungsprämissen. Der emanzipatorische Diskurs ist darauf ausgerichtet, dass das Gegenüber als Subjekt seines Handelns mit seinen tatsächlichen, eben auch sinnlich emotionalen Handlungsgründen verstanden und anerkannt wird. Prozesse der Offenlegung eigener subjektiver Wirklichkeiten, der Relativierung eigener Wahrnehmungen und Meinungen sowie der Reflexion im Sinne von Bewusstwerdungsprozessen über Lebensorientierungen werden durch metakommunikative Verständigung in Gang gesetzt.

# 2 Zur Entstehung des argumentativen Rückkoppelungsgesprächs

Die methodologische Grundbestimmung dieses Verfahrens geht von einem Aufgabenverständnis der Wissensbildung (Lorenzen & Schwemmer 1975) aus, welches nicht an der technischen Verwertbarkeit von Erkenntnissen orientiert ist, sondern auf die Gewinnung emanzipatorischen Reflexionswissens abzielt. Auf dieser Grundlage konstituieren sich Gegenstandsverständnis und Wissensideal der For-

schung, wobei unser Forschungsinteresse jenseits des sinnlosen Methodenstreits zwischen Vertretern quantitativer und qualitativer Verfahren liegt. Die oberflächliche (oberflächlich, weil abgekoppelt vom Gesamtzusammenhang des Forschungsgegenstandes) Auseinandersetzung mit der Frage Quantität vs. Qualität führte bekanntlich dazu, dass ernsthafte methodische Auseinandersetzungen über Forschungsziele und Erkenntnisinteresse(n) der jeweiligen Forschungsbemühungen weitgehend ausblieben. Positiv zu werten in der neueren Geschichte der Methodendebatte in der Psychologie sind zweifelsohne die Versuche einer Überwindung der Erklären-Verstehen-Kontroverse (Greve 1994, Dierstein 1995, Jüttemann 1991); diese waren vorwiegend darauf gerichtet, aus einer handlungstheoretischen Perspektive heraus die Intentionalität menschlichen Handelns zu erklären und mit Bezug auf die moderne analytische Wissenschaftstheorie den Erklärungsbegriff so weit zu fassen, dass er mit dem des Verstehens kongruent ist. Ein weiterer Meilenstein der Methodendebatte auf dem Weg zur Aufhebung der Erklären-Verstehen-Dichotomie ist die Berücksichtigung zentraler Prämissen der Cause-Reason-Debatte in der analytischen Philosophie und Handlungstheorie (Dray 1985, Beckermann 1985): Der Analyseschwerpunkt wird nicht nur auf die Sinndeutung einer Handlung (im Sinne des Verstehens) gelegt; subjektiv begründetes menschliches Handeln soll zunächst beschrieben und darüber hinaus explikativ erfasst werden. Bereits von Wright (1974) definierte Verstehen als logisches Verständnis von Handlungen und Handlungsmotiven, während Graumann (1984, S. 54f.) in der phänomenologischen Tradition argumentierend die "Rekonstruktion der Sinnstruktur situativen Sich-Verhaltens als methodische Aufgabe auch der Psychologie" fordert und Erklärung intentional auffasst.

Das argumentative Rückkoppelungsgespräch ist entstanden und konzeptionell weiterentwickelt worden im methodologischen Kontext zweier Denkrichtungen, die Wilhelm Kempf eingeschlagen und für die psychologische Methodenlehre ausbuchstabiert hat. Es handelt sich zum einen um den methodologischen Konstruktivismus und zum anderen um den Ansatz der subjektwissenschaftlich-interpretativen Psychologie.

# 2.1 Methodologischer Konstruktivismus

Der methodologische Konstruktivismus, dessen Grundprämissen in der konstruktiven Wissenschaftstheorie (Kamlah & Lorenzen 1967, Janich et al. 1974, Lorenzen & Schwemmer 1975, Schwemmer 1976) verankert sind, grenzt sich gegenüber einem sozialen Konstruktivismus ab, der davon ausgeht, dass jede Wirklichkeit lediglich eine subjektive, bestenfalls intersubjektiv geteilte Wirklichkeit darstelle, und postuliert im Unterschied dazu, dass subjektive, intersubjektive und transsubjektive Wirklichkeiten nicht gleichberechtigt nebeneinander existieren. Zwar stelle in der konstruktiven Wissenschaftstheorie jede Wirklichkeit eine menschliche Konstruktion dar, dennoch unterschieden sich die drei unterschiedlichen Grade von Wirklichkeit in der Art und Weise, wie die sie konstituierenden Aussagen begründet sind. Für die Wissenschaft ist diese Unterscheidung insofern folgenreich, als sie

unter der Transzendierung von Subjektivität den Anspruch einzulösen versucht, die Wirklichkeit in solchen Aussagen abzubilden, "deren Wahrheit anhand von explizit vereinbarten und methodisch begründeten Regeln ... nachgeprüft werden kann" (Kempf 2009b, S. 61).

Durch die Bezugnahme auf die Ebene der transsubjektiven Wirklichkeit wird das methodische Kriterium der Angemessenheit zur Überprüfung von Bedeutungszuweisungen und Prämissen menschlichen Handelns (ebd., S. 91) in den Vordergrund gerückt. Hiermit strebt der kempfsche Ansatz eine kritische Sozialforschung an, welche sich nicht auf die Konstitutionsbedingungen der subjektiven Welt des Menschen beschränkt, sondern darüber hinaus Orientierungshilfen für menschliches Dasein und Handeln zu bieten vermag. Kempfs methodologischer Konstruktivismus trägt daher nicht nur zur Überwindung der dem sozialen Konstruktivismus innewohnenden Widersprüchlichkeit bei der Handhabung des Wirklichkeitsbegriffs bei, sondern schafft darüber hinaus den erforderlichen methodologischen Überbau, durch welchen subjektwissenschaftliche Fragestellungen empirisch zugänglich gemacht und methodisch kontrolliert erforscht werden können.

### 2.2 Subjektwissenschaftlich-interpretative Psychologie

Erkenntnisgegenstand einer subjektwissenschaftlich-interpretativen Psychologie (Kempf 1992a, S. 90ff.) ist nicht die Beschreibung isolierter Tätigkeiten, sondern die Analyse der Handlungszusammenhänge und Kontexte, in denen diese Tätigkeiten stehen und aus welchen Personen ihre Handlungsvorsätze subjektiv als "logische" bzw. "vernünftige" Konsequenz ableiten. Der Frage nach der Vermittlung der objektiven sozialen und ökologischen Verhältnisse in subjektiven Handlungsprämissen kommt dabei eine Schlüsselrolle zu. Subjektwissenschaftlich-interpretative Psychologie versteht sich als empirische Wissenschaft und bedient sich des intentionalen Erklärungsmodells für die Aufstellung und empirische Überprüfung von Hypothesen über Prämissen menschlichen Handelns. Die in der qualitativen Forschung geforderte Offenheit ist stets gepaart mit Theoriegebundenheit, wobei Theorien hier eher den Charakter "universell anwendbarer methodischer Prinzipien" (ebd., S. 103) zur argumentativen Überprüfung der formulierten Interpretationen haben.

Aufgabe der Forschung ist es, über eine einfache Beschreibung individueller Subjektivität (als ersten Schritt des Interpretationsprozesses) hinauszugehen und zu analysieren, "wie individuelle Subjektivität gesellschaftlich hervorgebracht wird" (ebd., S. 105). Dies bedeutet zweierlei: Zum einen gilt, dass das Bewusstsein kein Abbild des Realen darstellt und dass das Reale nicht die willkürliche Konstruktion des Bewusstseins ist (Freire 1981, S. 51f.). Zum anderen ist davon auszugehen, dass "die Diskrepanz zwischen Sein und Bewußt-Sein nicht einfach auf einem Irrtum der Subjekte beruht, sondern auf gesellschaftlich bedingten Abwehrprozessen" (Kempf 1992a, S. 103). Die Rekonstruktion gesellschaftlich hervorgebrachter Subjektivität schließt das Problematisieren subjektiver Formen der Realitätsbewältigung (Rationalisierungen, Verdrängungen, Abwehrmechanismen etc.) und ihre

Funktion für den Einzelnen stets mit ein. Psychoanalytische Begriffe erfassen und verdeutlichen in diesem Zusammenhang Erfahrungen und Erlebnisstrukturen des Individuums sowie "Selbsttäuschungen in Richtung auf die meinem Handeln zugrundeliegenden Konflikte, widerstreitenden Interessen und Bedürfnisstrukturen etc." (Maiers 1994, S. 68).

Subjektwissenschaftlich-interpretative psychologische Forschung strebt verallgemeinerbare, über den konkreten Fall hinausgehende Handlungserklärungen an, wobei Verallgemeinern weder mit statistischer Repräsentativität assoziiert noch mit einem Wegabstrahieren der Subjektivität der Adressaten und deren Problemen verwechselt werden darf. Strukturelle Verallgemeinerung meint, typische, also theoretisch verallgemeinerte (Konflikt- oder Problem-)Konstellationen herauszuarbeiten, welche Handlungsmöglichkeiten, Handlungseinschränkungen und Handlungswidersprüche der Subjekte repräsentieren (vgl. Held 1989, S. 136; Baros & Reetz 2002).

### 2.3 Sozialpsychologische Rekonstruktion

Subjektwissenschaftlich-interpretative Forschung findet ihre konkrete empirische Umsetzung durch das textinterpretative Verfahren der sozialpsychologischen Rekonstruktion (Kempf 1991, S. 127ff., Kempf 2008, S. 37ff.). Die Berücksichtigung des dialektischen Gegensatzes zwischen Individuum und Gesellschaft sowie die Analyse von gesellschaftlich bedingten Abwehrprozessen, so wie sich diese in den Regeln der Wirklichkeitskonstruktion seitens der Individuen manifestieren, stehen im Mittelpunkt der Analyse. Davon ausgehend, dass Prozesse der Ideologisierung der Gesellschaft in Kommunikationsinhalten auf verschiedenen Ebenen ihre Entsprechung finden, wird der Fokus auf den latenten Kommunikationsinhalt gelegt, welcher die Perspektive bestimmt, unter der die Subjekte manifeste Kommunikationsinhalte interpretieren, d. h. bestimmte Aspekte aufnehmen, während sie andere (Fakten, Erfahrungen, Bedeutungen etc.) thematisch abwehren. In Anbetracht gegenstandsbedingter Differenzierungen kann die sozialpsychologische Rekonstruktion unterschiedliche Zielsetzungen bei der Textinterpretation verfolgen. Stehen zum Beispiel bei einer Interaktionsanalyse von Gruppendiskussionen (vgl. Kap. 4) oder konfrontativen Interviews Themenverschiebungen in dem zu analysierenden Text im Vordergrund, wird also Kommunikation als Materialisierung der Ausgestaltung von subjektiven Wirklichkeitskonstrukten aufgefasst; so steht bei der Textanalyse von partnerzentrierten Gesprächen die Rekonstruktion von subjektiven Handlungsprämissen im Vordergrund (Handlungsanalyse). Hierbei gehen wir davon aus, dass selbst Forschungsprozesse, die dezidiert auf einem Mitforscher-Verhältnis aller Beteiligten beruhen (etwa in der Handlungsforschung), nicht ohne Textinterpretation auskommen: Denn das, was als Gegenstand und Ergebnis eines nahezu symmetrischen Dialogs zwischen den Beteiligten präsentiert wird, kann nur in Textform vorliegen.

Die sozialpsychologische Rekonstruktion als Handlungsanalyse untersucht systematisch den latenten und den exterritorialisierten Gehalt eines Textes und stellt den Prozess der Textinterpretation vor folgende Aufgaben:

- Herausarbeitung der Regeln, nach denen die Subjekte ihre Wirklichkeit konstruieren.
- 2. Erhellung der Art und Weise, wie die Subjekte mithilfe dieser Konstrukte ihre Realität bewältigen.
- 3. Rekonstruktion jener Bedeutungsaspekte der Kommunikation, die von den Interviewten systematisch abgewehrt werden.
- Aufschlüsseln von Handlungsprämissen, unter denen für den Einzelnen das Ausblenden/Abwehren bestimmter Bedeutungsaspekte der Kommunikation (Handlungsvorsatz) subjektiv als "logische" bzw. "vernünftige" Konsequenz erscheint.

Zur Rekonstruktion dieser Handlungsprämissen (4.) notwendiger Analyseschritt:

5. Eruieren des sekundären Gewinns, welchen die Subjekte aus der Thematisierungsabwehr ziehen.

### 3 Die Konzeption des argumentativen Rückkoppelungsgesprächs

Sowohl bei der Untersuchung der abgewehrten Bedeutungsaspekte der Kommunikation (3) als auch bei der Herausarbeitung des sekundären Gewinns der Subjekte aus der Thematisierungsabwehr (5) ist die sozialpsychologische Rekonstruktion auf das systematische Einbeziehen externen Kontextes angewiesen. Insbesondere die angemessene Beantwortung der Frage nach dem sekundären Gewinn aus der Thematisierungsabwehr bedarf einer – über die Interpretationen des Interviewtextes bezüglich der aktuellen Forschungsthematik hinausreichenden – Kontexterweiterung, die auf eine umfassende Analyse der aktuellen Lebenspraxis der Subjekte gerichtet ist. Folgende Fragen sind zu untersuchen:

- (Inwieweit) Finden sich die im Kontext des latenten Gehalts der Kommunikation herausgearbeiteten Regeln der subjektiven Wirklichkeitskonstruktion auch in anderen Bereichen der aktuellen Lebenspraxis der Subjekte als Bestandteile ihrer Realitätsbewältigung wieder?
- Welchen Wert beziehen Subjekte aus einer Fixierung auf Handlungsmuster, die zwar kurzfristig einer vordergründigen, belastungslindernden Situationsbewältigung dienen, jedoch langfristig zu einer permanenten Reproduktion ihrer Problemlage führen (können)?

Konsequent zu Ende gedacht kann diese Analyse Aussagen ermöglichen über die *subjektive Welt* der Betroffenen, auf die sie sich in ihrem Handeln beziehen, sowie über die *unterschiedlichen Formen/Grade der Wirklichkeit*, aus denen sich diese ihre subjektive Welt konstituiert.

Die hierfür erforderliche Kontexterweiterung kann durch argumentative Rückkoppelungsgespräche ermöglicht werden, in denen neues Datenmaterial gewonnen wird, indem die im Laufe des gesamten Forschungsprozesses erarbeiteten Interpretationen zum Gegenstand eines Diskurses mit den Betroffenen gemacht werden. Es handelt sich um eine argumentative Geltungsbegründung von Interpretationen (Terhart 1981), bei der die Interviewten weder bevormundet werden noch eine reibungslose kommunikative Einigung angestrebt wird. Ziel ist vielmehr ein Beratungsprozess zwischen Interpreten und Interpretierten, welcher nach begründbaren Regeln abläuft (ebd., S. 786). Argumentation ist dabei in Anlehnung an Thiel (1980) zu verstehen als "eine Rede mit dem Ziel ..., die Zustimmung oder den Widerspruch wirklicher oder fiktiver Gesprächspartner zu einer Aussage oder Norm durch den schrittweisen und lückenlosen Rückgang auf bereits gemeinsam anerkannte Aussagen bzw. Normen zu erreichen".

Beabsichtigt ist die Herstellung einer möglichst symmetrischen Beziehung zwischen den am Forschungsprozess Beteiligten (Forschern und Betroffenen): Die anfänglich auf dem Informationsvorsprung des Forschers beruhende Asymmetrie wird in einem zweiten Schritt dadurch relativiert, dass beide Seiten in einen Argumentationsprozess im Sinne eines gemeinsamen Lernprozesses eintreten. Argumentieren ist hier eine "Redepraxis, bei der die Subjektivität der Beteiligten nicht das letzte Wort ist, sondern, genau wie alle anderen Inhalte und Voraussetzungen, zur Disposition gestellt wird" (Wohlrapp 2009, S. 44). Argumentieren unter dem "Prinzip der Transsubjektivität" (ebd., S. 479) meint zum einen eine reflexive Haltung gegenüber der eigenen Subjektivität. Zum anderen – und analog zur Forderung der anerkennenden Haltung dem Anderen gegenüber – geht es darum, dass sich der Interpret um Begründungen für seine Thesen bemüht, die seine Konstrukte für die Interviewten nachvollziehbar machen (ebd., S. 491).

Das argumentative Rückkoppelungsgespräch besteht grundsätzlich aus folgenden aufeinander aufbauenden Phasen:

- Festlegung des Interpretationsgegenstands: Ausgangspunkt sind Interpretationen von Themenbereichen, die für die Interviewten selbst Probleme und/oder Widersprüche in ihrer Lebenslage und Realitätsbewältigung darstellen. Die Beteiligten handeln ihr unterschiedliches Problemverständnis aus.
- 2. Ergebnisse der bisherigen Analyse den Betroffenen verfügbar machen: Bereits vorliegende Interpretationen werden den Subjekten offengelegt, und zwar nicht als (end-)gültige Resultate, sondern als Hypothesen, die nach einer Reihe von Analyseschritten und Überlegungen entstanden sind. Wichtig ist dabei, den Subjekten deutlich zu machen, dass es nicht darum geht, die Wahrheit zu finden, sondern Erlebnisweisen und Erlebniswelten auszutauschen und die eigenen Standpunkte zum Gegenstand der Kommunikation zu machen.
- Argumentieren und Fortsetzung des hermeneutischen Zirkels: Das Präsentieren von Analyseergebnissen dient als Anstoß für den eigentlichen Argumentationsprozess, in welchem alle Beteiligten die Möglichkeit haben, zu den Interpretationen des anderen argumentativ Stellung zu nehmen. Dieser Prozess zielt nicht

- vordergründig auf Konsensbildung ab; vielmehr sollen die Voraussetzungen für die argumentative Überprüfung auch unterschiedlicher Sichtweisen geschaffen werden.
- 4. Überwinden von Widerständen und Suche nach angemesseneren Formen der Situationsbewältigung: Die systematische Aufhebung der Trennung zwischen Interpretation und Intervention ermöglicht es, bereits im Forschungsverlauf einen Perspektivenwechsel im Sinne von Prämissenerweiterung für alle am Forschungsprozess Beteiligten herbeizuführen. Auf der neuen Prämissengrundlage dürften sich Meinungsbildungsprozesse verändern, d. h. Forschungssubjekte stellen Behauptungen über mögliche Wirkungen von Handlungen auf und entwickeln ein anderes Verständnis für die Erforderlichkeit der Ausführung von Handlungen, damit bestimmte Sachverhalte eintreten.

Diese methodische Konzeption positioniert sich gegenüber den unterschiedlichen Ebenen der Verständnisbildung neu: Im Feld der intersubjektiven Verständnisbildung lässt der Interpret die Interviewten an seinen Bedeutungsbeschreibungen und Interpretationen teilhaben. Dies ist prinzipiell auch bei der kommunikativen Validierung der Fall: Kommunikative Validierung versucht, eine Einigung zwischen dem Forscher und den Betroffenen hinsichtlich der Interpretation der Untersuchungsergebnisse zu erreichen (Klüver 1979). Im Extremfall wird sogar dafür plädiert, dass sich nur jene Interpretationen behaupten können, die sowohl in einer Gemeinschaft wissenschaftlicher Interpreten als auch bei ihrer Reflexion mit den Betroffenen selbst nicht widerlegt werden konnten (Heinze & Klusemann 1979). Sozialforschung wäre jedoch überflüssig, könne oder dürfe sie nicht über das hinausgehen, was die handelnden Subjekte selbst schon von sich bzw. ihrem gesellschaftlichen Umfeld wüssten bzw. zu wissen meinten. Interpretationen könnten durchaus auch dann Gültigkeit besitzen, wenn die Betroffenen sie ablehnten (Moser 1995, S. 67). Interpretationen können aber auch dann ungültig sein, wenn die Betroffenen ihnen – um in möglichst gutem Lichte zu erscheinen – zugestimmt haben (Kempf 2003a, S. 33). Nicht zuletzt stellt sich die grundlegende Frage, inwieweit der im Konzept der kommunikativen Validierung postulierte Dialog mit den Betroffenen zum Zwecke der Objektivierung von Untersuchungsergebnissen tatsächlich praktiziert wird, wenn die Rolle des handelnden Subjekts lediglich auf die eines relevanten Datenlieferanten für die Absicherung der Interpretationen reduziert wird.

Während kommunikative Validierung auf der Ebene der intersubjektiven Verständigung stehen bleibt und dabei mehr oder minder unterstellt wird, beide subjektive Wirklichkeiten hätten den selben Geltungsanspruch bzw. könnten gleichberechtigt bestehen bleiben, nimmt das argumentative Rückkoppelungsgespräch genau dort eine besondere Stellung ein: Im Diskurs mit den Betroffenen entstehen Dissenssituationen, in denen erst unterschiedliche Wahrnehmungen der Wirklichkeit erfahrbar bzw. bewusstseinsfähig gemacht werden (vgl. auch Hartmann & Werbik 2001). Ausgehandelt werden ferner Meinungen, d. h. sprachlich formulierte Erfahrungen als Behauptungen über die Wirkungen von Handlungen bzw. als Behaup

tungen über die Erforderlichkeit der Ausführung von Handlungen für das Auftreten von bestimmten Sachverhalten (Lorenzen & Schwemmer 1975, S. 154).

Das Feld der transsubjektiven Verständnisbildung betreten Interpret und Interpretierte im Verlauf eines argumentativen Rückkoppelungsgesprächs gemeinsam. Dabei stehen unterschiedliche Beschreibungen sozialer und/oder subjektiver Bedeutungswelten, Erklärungen ihrer Entstehungs- und Konstitutionsbedingungen sowie Analysen der Interaktion von Wirklichkeitskonstrukten nicht gleichberechtigt nebeneinander. Für die Überprüfung der Angemessenheit von Bedeutungszuweisungen wird eine relationale Perspektive eingenommen, aus der die Orientierungsfunktion untersucht wird, welche Bedeutungen für das menschliche Handeln besitzen (Kempf 2009b, S. 91). Das gemeinsame Betreten des Feldes der transsubjektiven Verständnisbildung zielt darauf ab, auf Grundlage des transsubjektiv begründeten Wissens Reflexionen im Sinne von Bewusstwerdungsprozessen über Lebensorientierungen in Gang zu setzen. Da letztere an der Konstitution von Handlungsprämissen maßgeblich beteiligt sind und dadurch die Art und Weise bestimmen, wie Individuen die objektseitig definierten Randbedingungen (unbewusst) interpretieren (Kempf 2003a, S. 259), kann deren Bewusstmachung auf dieser erweiterten Analyseebene methodisch dazu dienen, Sinnzusammenhänge des Handelns und Erlebens zu rekonstruieren und mithin subjektive Gründe für die Exterritorialisierung von bestimmten Inhalten aus dem subjektiven Bewusstseinshorizont in der Kommunikation zu ergründen.

### 4 Empirische Anwendungen

# 4.1 Untersuchungsfelder

Das argumentative Rückkopplungsgespräch wurde bisher vorwiegend in der interkulturellen Migrations- und Bildungsforschung angewendet. Das Spektrum der bearbeiteten Fragestellungen ist breit gefächert: Es reicht von (a) Analysen adoleszenter Generationenbeziehungen (vgl. Baros 2001) und des subjektiven Integrationsverständnis von Migrant(inn)en (Baros & Reetz 2002) über (b) Studien zu informellen sozialen Netzwerken älterer Migranten (Baros 2010) sowie prekären Aspekten des Daseins in der Migration (Baros & Manafi 2009) bis hin zu (c) Untersuchungen der Konfliktkommunikation zwischen Migranteneltern und der deutschen Schule (Reetz 2007) sowie zur (d) Erforschung der Antwortmuster von Rezipienten in Projekten der Medienbildung (Baros 2012, i.E.). Allen Projekten gemeinsam ist die Hinzuziehung externen Kontextes, um methodisch kontrollierte Textinterpretationen zu ermöglichen. Der methodischen Maxime folgend, wonach Geschichten aus mindestens drei Perspektiven erzählt werden können (erzählendes Subjekt, Interaktionspartner, Beobachter) (Werbik & Appelsmeyer 1999), wurden mehrere Erzählperspektiven berücksichtigt: Adoleszente und deren Eltern über Konflikte in der Familie (a); Texte aus unterschiedlichen (interkulturellen) Interviewer-Interviewte-Konstellationen (b); Lehrer(innen), Eltern und Schüler(innen) über Ursachen von Schulversagen (c) sowie Gruppendiskussionen mit Textproduzenten als Repräsentanten latenter Klassen von Antwortmustern. Die gleichzeitige Explikation dieser Perspektiven unter Hinzuziehung theoretischen Wissens bildete eine unerlässliche Voraussetzung für den Interpretationsprozess. Im Folgenden wird exemplarisch auf zwei Anwendungsgebiete des argumentativen Rückkoppelungsgesprächs eingegangen, die jeweils unterschiedliche Zielsetzungen und methodische Schwerpunktsetzungen des Verfahrens markieren. Zum einen handelt es sich um die Fortsetzung des hermeneutischen Zirkels im Rahmen einer *Handlungsanalyse*, wobei die (sozialpsychologische) Rekonstruktion der Handlungsprämissen der Befragten im Mittelpunkt steht. Zum anderen geht es um *Interaktionsanalyse* von Gruppendiskussionen über Untersuchungsergebnisse einer quantitativen Inhaltsanalyse (LCA). Dabei steht die Modifikation eines Themas durch die Interaktion der Beteiligten miteinander im Vordergrund.

#### 4.2 Handlungsanalyse

In der kommunikations- und handlungstheoretischen Analyse der Elternhaus-Schule-Interaktion (Reetz 2007) wurde der Frage nachgegangen, wie die Beteiligten (Migranteneltern, Schüler und Lehrer) den schulischen Misserfolg wahrnehmen und bewältigen und wie sich die Konfliktkommunikation zwischen Elternhaus und Schule gestaltet. Unter Bezugnahme auf Handlungs- und Konflikttheorien (Kempf 1993, Pruitt & Rubin 1986) sowie die Theorie der informativen Distanz (Oevermann 1974) konnten Entstehungsbedingungen divergierender Erwartungen, die beide Seiten aneinander haben, verdeutlicht werden. Ebenso konnte nachgewiesen werden, dass sich Konfliktanlässe aus der fehlenden Verzahnung der beiden Erziehungsinstanzen als Folge einer institutionell bedingten informativen Distanz ergeben. Unzureichende Kommunikation führt – so die unmittelbare Folge – zu Informationsdefiziten, die durch die Übertragung von tradiertem Wissen verstärkt werden. Da die damit verbundenen wechselseitig generalisierten Erwartungen nicht hinterfragt werden, gehen beide Seiten selbstverständlich davon aus, dass ein gemeinsamer Deutungshorizont des Aufgaben- und Rollenverständnisses vorliegt. Auf der Grundlage dieses gegenseitigen Vertrauens kommt es oft zu Scheinübereinstimmungen. Dieser auf verdeckter Perspektivendivergenz beruhende Scheinkonsens bestimmt die Kommunikation der Beteiligten. Sie kennen die von den eigenen unterschiedenen Erwartungen des anderen nicht und sind sich zudem infolge ihrer Perspektivendivergenz der tatsächlichen Unvereinbarkeit ihrer wechselseitigen Erwartungen nicht bewusst. Auf Grundlage dieses Wissens wurde in Rückkoppelungsgesprächen mit den Eltern (Reetz 2007, S. 120ff.) das Feld der transsubjektiven Verständnisbildung betreten: Die unterschiedlichen Sichtweisen aller Beteiligten wurden argumentativ überprüft und zur Disposition gestellt. Die Eltern wurden mit nicht-intendierten Folgen ihres Handelns konfrontiert, damit sie die gesellschaftliche Vermitteltheit ihrer Probleme als Konfliktpotenzial erkennen konnten: Vermeintliche Rückbesinnungstendenzen auf traditionelle Erziehungspraktiken konnten unter ihren aktuellen Lebensumständen als Momente und Ausdrucksformen der individuellen Ausgestaltung ihrer Migrationserfahrung verstanden werden. Die Erkenntnis, dass die Förderung der Selbständigkeit und die Unterstützung der Autonomiebestrebungen der Kinder einen Kontrollverlust über die kindlichen Lebensentwürfe und Zukunftspläne bedeuten könnte, wurde als sekundärer Gewinn herausgearbeitet, den die Eltern aus dem Praktizieren rigider Erziehungsmethoden ziehen konnten. Als spezifisches Moment einer durchgängigen Thematisierungsabwehr konnte die versuchte Aufrechterhaltung der verdeckten Perspektivendivergenz zwischen den Eltern und der Klassenlehrerin im Gespräch identifiziert werden, was im Kontext der bisherigen Migrationsbiografie der Eltern analysiert wurde. Das argumentative Rückkoppelungsgespräch bietet ferner die Möglichkeit, das schulische Problem in Form eines *gemeinsamen Definitionsprozesses* mit den Beteiligten zu analysieren. Die unterschiedlichen Perspektiven werden im Dialog offengelegt und argumentationszugänglich gemacht, mit dem Ziel, auf einem höheren Erkenntnisniveau eine gemeinsame Problemlösungsstrategie zu entwickeln.

#### 4.3 Interaktionsanalyse

Bausteine des argumentativen Rückkoppelungsgesprächs wurden auch in Gruppendiskussionen eingesetzt, um eine *klassenbezogene Typikalität* von Ergebnissen der inhaltsanalytischen Anwendung der Latent-Class-Analyse (LCA) (Kempf 2008, S. 119f.) erarbeiten zu können. Im Rezipientenforschungsprojekt Globalisierung, Nationalismus und die Medien (Baros 2012, i.E.) wurde dies dadurch erreicht, dass Textproduzenten selbst als Repräsentanten charakteristischer Antwortmuster (d. h. latenter Klassen) an Gruppendiskussionen teilnahmen. Die mittels der LCA identifizierten Klassen beruhen auf der quantitativen Inhaltsanalyse von Texten (n = 108), welche von Lehramtstudierenden und Lehrer(inne)n in Griechenland geschrieben worden waren. In diesen Texten geben sie ihren subjektiven Eindruck von einem Lied (bezogen auf Inhalt, Bild und Ton/Melodie) wieder. Einige der wichtigsten inhaltlichen Botschaften der Rezipiententexte sind Globalisierungskritik, Bedeutungsverlust des Nationalstaates, wirtschaftliche/soziale Verelendung im eigenen Land aufgrund fremden Einflusses, zurück zu den eigenen kulturellen Wurzeln als ultima ratio, die – gepaart mit subtil transportierten nationalistischen Deutungsmustern – vom Liedtext vermittelt werden. Die LCA identifizierte drei Klassen und brachte eine interessante Erkenntnis: Je deutlicher die Nationalismusgefahr in den Inhalten des Liedes erkannt wurde, desto seltener argumentieren die Probanden in der Wir-Form (z. B. "Wir müssen gegen die Globalisierung ankämpfen"; "Wir sind angesprochen und müssen dagegen etwas unternehmen" etc.), desto seltener gehen sie auf die im Lied angesprochenen Probleme und Konflikte im Einzelnen ein (Klasse 3). Komplementär dazu ist das Profil von Klasse 1: Je mehr die Probanden den Inhalten zustimmen, desto häufiger argumentieren sie in der Wir-Form und desto öfter heben sie die Bedeutung der national-kulturellen Identität als wichtigen Gegenpol zur fremdbestimmten Homogenisierung hervor. Und schließlich: Probanden, die Klasse 2 angehören, erkannten in der Tendenz nationalistische Inhalte des Liedtextes oft; selbst dann, wenn sie die dort thematisierten Probleme (z. B. Folgen der Globalisierung, wirtschaftliche und kulturelle Fremdbestimmung, Bedeutungsverlust des Nationalstaates) relativ ernst nehmen.

Für die Gruppendiskussionen (Schäffer 2006) wurden drei Gruppen mit jeweils drei Personen zusammengestellt, die nach der (Membership-)Wahrscheinlichkeit je einer der drei Klassen angehören. Insgesamt wurden vier Gruppendiskussionen geführt; eine mit jeder Dreiergruppe und eine mit der aus neun Personen bestehenden Gesamtgruppe. Das Ziel der Gruppendiskussionen bestand zunächst darin, herauszufinden, wie sich die Probanden zur Globalisierung und deren Folgen positionieren und welche Bedeutung sie Inhalten wie Anerkennung von Diversität, nationaler und kultureller Identität, Politik, Patriotismus etc. zuweisen – Themen, die explizit oder implizit im Lied vorkommen. Jede Gruppendiskussion verlief in zwei aufeinanderfolgenden Phasen. In der ersten Phase wurde über das Lied diskutiert, über die mögliche Intention des Produzenten, über die angesprochene Personengruppe – ein Austausch also über den persönlichen Gesamteindruck (intersubjektive Verständnisbildung). Die Konfrontation der Gesprächspartner mit den Ergebnissen der LCA erfolgte erst in der zweiten Phase (transsubjektive Verständnisbildung), deren Ziel darin bestand, das Erleben von Fremd- und Selbstzuordnung argumentativ zu bearbeiten, sich zu den eigenen individuellen Deutungsmustern zu verhalten, Prozesse der Selbstidentifikation bzw. -disidentifikation in Gang zu setzen, ja die eigene Identität auszuhandeln. In der sich daran anschließenden Diskussion, an der alle drei Gruppen teilnahmen, ging es um die Schaffung eines Orts zur argumentativen Auseinandersetzung mit individuellen Positionierungen und Handlungsorientierungen; um - auch wenn nur in diesem quasi-experimentellen Setting – die Herstellung von Kanälen zur Ermöglichung eines agonistischen Pluralismus (Mouffe 2008) zur Artikulation von unterschiedlichen Interessenlagen und subjektiven Lebensinteressen.

Zusammenfassend lässt sich festhalten: Konstituiert sich transsubjektives Wissen in der Handlungsanalyse durch die fallbezogene Überprüfung der Angemessenheit von subjektiven Bedeutungszuweisungen und Handlungsprämissen, so ist dieses Wissen am Beispiel der konkreten Interaktionsanalyse auf klassifizierte Antwortmuster bezogen. Diese Antwortmuster deuten ihrerseits auf typische, theoretisch explizierbare Formen der Manifestation von Prozessen der Ideologisierung in der Gesellschaft in Inhalten der Kommunikation hin. Die Konfrontation der Befragten mit den Antwortklassen und deren charakteristischen Ausprägung bedeutet eine Erweiterung des für Interaktionsanalysen erforderlichen Makro-Kontextes der Kommunikation: Es wird ein neuer objektseitig definierter Rahmen bereitgestellt, zu welchem sich die Gruppendiskussionsteilnehmer – wie auch immer – verhalten, ihre eigenen subjektiven Sinngebungsprozesse und Handlungsorientierungen Preis geben bzw. zur Disposition stellen. Die oben kurz skizzierte Vorgehensweise bei den konfrontativen Gruppendiskussionen markiert nicht zuletzt auch einen methodischen Zugewinn in Richtung auf Integration quantitativer und qualitativer Verfahren – eine Integration, die das Ziel verfolgt, wissenschaftliche Erkenntnis kontinuierlich auf- und auszubauen und darüber hinaus Reflexionsprozesse über Sinnzusammenhänge menschlichen Handelns und Erlebens in Gang zu setzen.

# Die Vision einer integrativen Psychologie im Spannungsfeld zwischen Natur- und Kulturwissenschaft

Markus Kiefer & Klaus Hönig

"Erst wenn wir keine dieser Geschichten mehr ausblenden oder als Grauzone behandeln, mit der sich die Psychologie nicht zu beschäftigen braucht, können wir dem Gegenstand der Psychologie gerecht werden. Oder anders ausgedrückt: Eine gegenstandsangemessene Psychologie ist nur in der Einheit von Natur- und Kulturwissenschaft denkbar."

Kempf 2009a, S. 37f.

Menschen erleben sich als erkennende, fühlende und handelnde Individuen, deren Existenz in einen kulturellen und historischen Kontext eingebettet ist. Die subjektive Sicht auf die Welt, geprägt durch ein komplexes Gefüge von physikalischen, biologischen und sozialen Faktoren, ist ein wesentlicher Bestandteil des Selbstverständnisses und der Identität des Menschen. Gleichzeitig ist diese subjektive Welt (Rohracher 1963) Basis seines Handelns, da in ihr die subjektseitigen Gründe seines Verhaltens verankert sind (Kempf 2009a, Kempf & Kiefer 2009b).

Menschliches Handeln hat aufgrund der Herausbildung einer Symbolsprache eine vergleichsweise hohe Flexibilität erreicht. Die damit notwendigerweise einhergehende Bedeutungsgebundenheit menschlichen Handelns macht es aber auch offen für Interpretationen, da keine eindeutigen Entsprechungen zwischen objektseitig definierten Umweltgegebenheiten und der bisweilen sehr komplexen subjektseitig definierten Situation bestehen (Hönig 2009). Weil diese Bedeutungen den "Gegenständen" unserer inneren und äußeren Welt nicht einfach innewohnen, sondern vielmehr das Ergebnis biografischer Erfahrungen, sozialer Austauschprozesse und kultureller Traditionen darstellen, ist ihre Konstitution in hohem Maße perspektiven- und kulturabhängig (Kosslyn 2005).

Die wissenschaftliche Psychologie hat zum Ziel, menschliches Erleben und Verhalten zu beschreiben und zu erklären. Die Erklärung menschlichen Erlebens und Verhaltens aufgrund psychologischer, mentaler Prozesse ist das genuine Anliegen einer sich als nomothetisch verstehenden Psychologie, die allgemeine Theorien über psychische Phänomene aufzustellen versucht. Unserer Ansicht nach dient aber auch der hermeneutische Ansatz einer verstehenden Psychologie letztendlich der Erklärung menschlichen Erlebens und Verhaltens (vgl. auch Albert 1991), so dass

wir in dem vorliegenden Text nicht weiter zwischen einer erklärenden und einer verstehenden Psychologie differenzieren. Im Zuge psychologischer Forschungstätigkeit steht die oben beschriebene Subjektivität des Gegenstandes in einem Spannungsfeld mit den strengen wissenschaftstheoretischen Kriterien, die an eine Erfahrungswissenschaft herangetragen werden (Eimer 1996), Menschliches Fühlen, Denken, Wissen und Wollen ist nur aus der Erste-Person-Perspektive erfahrbar und damit notwendigerweise subjektiv: Ein "Ich" fühlt, denkt, weiß und will etwas (Prinz 1996). Diese subjektive Welt des Menschen ist Gegenstand eines kulturwissenschaftlichen Ansatzes in der Psychologie. Gleichzeitig ist die Psychologie als Erfahrungswissenschaft den Kriterien der Objektivität und der Falsifizierbarkeit von theoretischen Aussagen verpflichtet. Diese wissenschaftstheoretischen Kriterien erfordern die Einnahme einer Dritte-Person-Perspektive und die Beschränkung auf objektivierbare Sachverhalte (Eimer 1996, Velmans 1991). Die Fokussierung auf die Dritte-Person-Perspektive erfolgt am konsequentesten in der naturwissenschaftlich ausgerichteten experimentellen bzw. neurowissenschaftlichen Psychologie.

Um dem Gegenstand der Psychologie gerecht werden zu können, müssen nicht nur beide Wege eingeschlagen werden, sondern auch die Forschungsergebnisse beider Herangehensweisen aufeinander bezogen werden (Hönig 2009). Diese reziproke Beziehung setzt aber voraus, dass die Unterschiede der beiden Pole (hinsichtlich Gegenstandsverständnis und Wissensideal) adäguat berücksichtigt werden und ein gegenseitiges Bewusstsein geschaffen wird für die methodischen und methodologischen Probleme der jeweiligen Paradigmen. Denn natur- und kulturwissenschaftliche Paradigmen der Psychologie unterscheiden sich sowohl im Gegenstandsbereich (objektivierbare Sachverhalte vs. subjektive Erfahrungen) als auch in den wissenschaftlich akzeptierten Methodologien und Methoden zur Gewinnung von Erkenntnissen über psychische Phänomene (experimentell quantitativ vs. subjektbezogen qualitativ). Eine erfolgreiche Konvergenz von kultur- und naturwissenschaftlichen Paradigmen kann aber vermutlich nur einer Transdisziplinarität gelingen, welche die subjektive Welt des Menschen als sinn- und bedeutungsstrukturierende Wirklichkeit in das naturwissenschaftliche Paradigma mit einbezieht (Libet 2005).

Der vorliegende Beitrag will zeigen, dass die Notwendigkeit einer solchen Konvergenz seit Anbeginn der akademischen Psychologie von deren Wegbereitern immer wieder erörtert wurde, aber erst in jüngster Zeit durch das Aufkommen der kognitiven und sozialen Neurowissenschaften einen ersten systematischen empirischen Widerhall findet. Er endet mit der Skizze der Vision einer integrativen Psychologie, welche die klassischen Grenzen zwischen Natur- und Kulturwissenschaft überschreitet.

# 1 Integrative Ansätze in der wissenschaftlichen Psychologie: Ein historischer Abriss

# 1.1 Wundt: Integratives Modell einer historischen Kultur- und Sozialpsychologie

Eine Konvergenz natur- und kulturwissenschaftlicher Paradigmen ist sowohl theoretisch als auch forschungsprogrammatisch bereits in den Arbeiten des Gründungsvaters der akademischen Psychologie, Wilhelm Maximilian Wundt (1832-1920), angelegt. Es gilt als dessen unbestrittenes Verdienst, die Psychologie als wissenschaftliche Disziplin etabliert und ihr zu weltweiter Anerkennung verholfen zu haben. Die allfällige Anerkennung für Wundt als prototypischen Experimentator und unbeugsamen Vorkämpfer für eine streng naturwissenschaftliche Psychologie (Wundt 1874) lässt allerdings zwei weitere untrennbare Bestandteile seines integrativen Modells sträflich außer Acht: nämlich seine konsequent antimaterialistisch inspirierte geisteswissenschaftliche Fundierung der Psychologie sowie seinen Aufbau einer historischen Kultur- und Sozialpsychologie (Wundt 1912).

#### 1.2 Lurija: Die Zweieinigkeit der Neuropsychologie

Alexander Romanowitsch Lurija (1902-1977) gilt heute insbesondere als Begründer der Neuropsychologie. Der Name der Disziplin ist bereits Programm, und zwar insofern, als er auf eine Annäherung zwischen dem Neuralen und dem Psychischen drängt. Für den Schüler und Freund des einflussreichen russischen Psychologen Lew S. Wygotski (1896-1934) bildeten Psychologie, Neurologie und Sprachforschung naturgemäß die Hauptarbeitsgebiete. Die Komplexität und Fruchtbarkeit seiner Sicht der Wissenschaft und der Welt ist nicht zuletzt der Tatsache geschuldet, dass sich sein Schaffen konsequent im Spannungsfeld zwischen "klassischer" und "romantischer" Wissenschaft bewegte und damit den Konflikt zwischen der Suche nach allgemeingültigen Gesetzen (meist Naturgesetzen; Nomothetik) und der Analyse zeitlich und räumlich einzigartiger Wissensgegenstände (Idiographie) aushält und aufzulösen versucht.

Das Aushalten dieses Konflikts wird geradezu in der Verschiedenheit seiner Schriften reflektiert: Auf der klassischen Seite die systematisch-wissenschaftlicheren Werke (Luria 1966a, 1966b, 1970), allen voran sein monumentaler Klassiker *Die höheren kortikalen Funktionen des Menschen*, und auf der anderen Seite die eher novellistisch gehaltenen romantischen (Luria 1968, Lurija 1991). Für Lurija bildeten das Klassische und das Romantische zwei Seiten einer Medaille, deren gemeinsame Verfolgung unter feinfühliger Betrachtung der Gesamtheit der Daseinsbedingungen eines Menschen mehr hervorbringen kann als die der Summe ihrer Teile. Nach Lurija beschäftigt sich die Neuropsychologie mit den höheren kortikalen Funktionen und den kognitiven Tätigkeiten des Menschen und wurzelt somit gleichsam im Biologischen und im Sozialen, in Natur und Kultur. Erst die Synthese beider – im Sinne einer Versöhnung des Objektiven mit dem Subjektiven, des Klas-

sischen mit dem Romantischen, des Erklärenden mit dem Beschreibenden – erlaube die Überwindung der "Krise der Psychologie" (Vygotsky,1927).

#### 1.3 Hebb: Neuronale Netze, Genetik und Sozialisation

Als Donald Olding Hebb (1904-1985) 1949 seine monumentale Monografie *The Organization of Behavior* veröffentlichte, war noch kaum bekannt, auf welche Weise die Struktur und Organisation des zentralen Nervensystems jene Funktionen hervorbringen, die in psychologischen Studien beobachtet werden. Weil die damals verfügbaren Methoden noch keine empirisch gut begründeten Aussagen über die reale Funktion und Struktur des menschlichen Nervensystems erlaubten, bezogen sich Hebbs Ideen auf ein konzeptuelles Nervensystem. Er war der Überzeugung, dass befriedigende Antworten auf psychologische Fragestellungen nur dann zu erwarten seien, wenn die verschiedenen Analyseebenen der Psychologie als sich gegenseitig befruchtende berücksichtigt werden. Im Hinblick auf ihre wissenschaftliche Gültigkeit seien diese Ebenen als gleichberechtigt anzusehen. Erst deren Verbindung erlaube ein höheres Maß an Verständnis und Vorhersage.

Die grundlegende Vorstellung für eine solche Integration liege in dem Verständnis, auf welche Weise die neuronalen Netzwerke, die unsere Gedanken, Gefühle und Handlungen realisieren, sowohl durch die Gene als auch die Erfahrung geformt werden. In der Verbindung von neuronalen Netzen, Genetik und Sozialisation sah Hebb einen allgemeinen Zugang zu sämtlichen Aspekten menschlicher Kognition und Emotion. Die Vereinigung sozialer, kultureller, differentieller, experimenteller und physiologischer Aspekte bilde die Grundlage der Psychologie, von der aus letzten Endes auch die verschiedenen Formen von pathologischem Verhalten und Erleben erst befriedigend verstehbar werden.

# 2 Kognitive und soziale Neurowissenschaften

Angesichts der enormen Fortschritte in dem Bemühen, dem menschlichen Gehirn, dem biologischen Substrat von Denken, Fühlen und Handeln, seine Geheimnisse zu entreißen, sind neurowissenschaftliche Methoden für die Erforschung psychischer Phänomene kaum noch wegzudenken. Neurowissenschaftliche Techniken der funktionellen Bildgebung machen psychische Vorgänge anhand physiologischer Prozesse im Gehirn sichtbar, unmittelbar während wir denken, fühlen und handeln. Selbst flüchtige Gedanken sind durch die Messung der Hirnaktivität materiell fassbar und scheinen dadurch objektivierbar. Neurowissenschaftliche Techniken erlauben es uns, dem Gehirn gleichsam bei der Arbeit zuzuschauen und öffnen ein bedeutsames Fenster zu unserem Geist (Kiefer 2009).

Die Erforschung der biologischen Korrelate kognitiver und sozialer Prozesse ist zwar bereits seit einigen Jahrzehnten im Gang (Cacioppo & Berntson 1992), hat mit dem Einsetzen der funktionellen Hirnbildgebung jedoch einen rasanten Entwicklungsverlauf genommen; mit der Konsequenz, dass sich die in Kapitel 1 skizzierten his-

torischen Integrationsideen noch nie so nahe an ihrer empirischen Realisierung befanden wie heute. Spätestens 2001 etablierte sich ein viel versprechendes Forschungsfeld, das die Werkzeuge der kognitiven Neurowissenschaft mit Fragestellungen und Theorien aus verschiedenen sozialwissenschaftlichen Bereichen wie der Sozialpsychologie, der Wirtschafts- und Politikwissenschaft verbindet.

In Kapitel 2.1 wird die Fruchtbarkeit dieses integrativen Vorgehens anhand einiger empirischer Untersuchungen aus den Bereichen Lernen und Gedächtnis exemplarisch illustriert. Die Beispiele sollen aufzeigen, wie ein konzertiertes Vorgehen biologischer, psychologischer und sozialwissenschaftlicher Forschung zu Erkenntnissen über die Natur menschlichen Verhaltens und Erlebens führen kann, die der biologischen *und* der gesellschaftlich-kulturellen Bedingtheit des Menschen Rechnung tragen (für eine weiterführende Diskussion vgl. Hönig 2009).

#### 2.1 Lernen und Gedächtnis

Begriffliches Wissen über die Bedeutung von Gegenständen, Vorgängen und abstrakten Ideen ist im semantischen Gedächtnis, einer Teilstruktur des Langzeitgedächtnisses, repräsentiert. Das semantische Gedächtnis spielt eine wichtige Rolle bei der Identifikation von Gegenständen, objektorientiertem Handeln sowie beim Sprachverständnis und in der Sprachproduktion (Herrmann 1985, Herrmann & Grabowski 1994, Kiefer 1999).

Mittlerweile gilt es als gut etabliert, dass die Summe unserer Erfahrungen mit der gegenständlichen Welt kategorial durch Begriffe repräsentiert wird (belebt vs. unbelebt, Werkzeuge vs. Tiere oder Pflanzen etc.), die im semantischen Gedächtnis gespeichert werden (Hoffmann 1986, Kiefer 1999). Wie unser begriffliches Wissen nun genau im Gehirn repräsentiert wird, ist Gegenstand einer anhaltend heftigen Diskussion. Traditionellerweise werden unsere Begriffe sowohl in der Kognitionswissenschaft, der Neurowissenschaft als auch im philosophischen Diskurs als situationsinvariante mentale Wissensentitäten aufgefasst (Stabilität von Begriffen), die eine amodale Sprache des Geistes konstituieren, und zwar unabhängig von Wahrnehmung und Handlung (Anderson 1983, Anderson & Bower 1973, Fodor 2001, Tyler & Moss 2001).

Neuerdings jedoch wird die Idee amodaler Begriffsrepräsentationen zunehmend in Frage gestellt (Kiefer 2008): Stattdessen konvergieren die empirischen Befunde aus Verhaltensuntersuchungen, neuropsychologischen Beobachtungen, elektrophysiologischen Aufzeichnungen und Untersuchungen mit räumlich hoch auflösenden bildgebenden Verfahren (funktionale Magnetresonanz-Tomographie [fMRT], Positronen-Emissions-Tomographie [PET]) zusehends auf eine Stützung modalitätsspezifischer semantischer Gedächtnissysteme (Barsalou et al 2003, Humphreys & Forde 2001, Kiefer & Spitzer 2001, Pulvermüller 2005). Dabei wird in Analogie zu den wohl etablierten sensorischen und motorischen neuronalen Karten angenommen, dass begriffliche Objektmerkmale (z. B. visuelle, akustische, handlungsrelevante) ebenfalls "leibnah" im Gehirn repräsentiert sind (embodied cognition; Gallese & Lakoff 2005, Lakoff & Johnson 1999), und zwar insofern, als sie in

modalitätsspezifischen semantischen Karten in sensorischen und motorischen Hirnregionen kodiert werden (Kiefer et al. 2007).

#### 2.1.1 Erfahrungsabhängige Plastizität begrifflicher Repräsentationen

Will man jedoch den Einfluss der konkreten Erfahrung mit den Dingen in der Welt auf die Begriffsbildung untersuchen, so sieht man sich unweigerlich dem Problem gegenüberstehen, dass unser begriffliches Wissen über unsere gegenständliche Welt im Erwachsenenalter bereits hochgradig überlernt ist. Andererseits muss angenommen werden, dass unsere Sinneserfahrungen, die wir beim Begriffserwerb mit verschiedenen Gegenständen gemacht haben, in Umfang und Intensität sehr stark interindividuell variieren. Um diese Schwierigkeiten zu umgehen, ließen Kiefer und Kollegen (Kiefer et al. 2007) ihre Probanden neuartige Objekte ("Nobjects") erlernen (vgl. Abb. 1). In den Trainingssitzungen wurden die Namen eines jeden visuell präsentierten Nobjects gelernt sowie dessen Zugehörigkeit zu zwei Arten von Objektkategorien: Objekte der Kategorien Umrissform waren ausschließlich über ihr weitgehend identisches äußeres Erscheinungsbild definiert; Objekte der Kategorien funktionales Merkmal über ein bestimmtes hervorstechendes Merkmal, das zu einer spezifischen Handlung auffordert. Im letzteren Fall unterschieden sich die Objekte in der Umrissform.

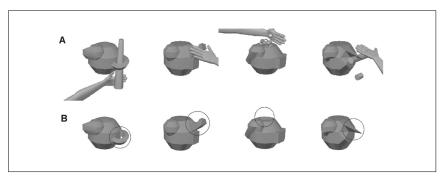


Abbildung 1: Künstliche Objekte (Nobjects) und Trainingsgruppen: (A) Handlungspantomime, (B) Zeigebewegung (nach Kiefer et al. 2007)

Ein wichtiger Aspekt der Studie war die Einteilung der Probanden in zwei Gruppen: Während die eine Gruppe (Handlungspantomime) bedeutungsvolle Handlungen bezüglich des Detailmerkmals ausführte, sollte die andere (Zeige-Gestik) lediglich mit dem Finger auf Merkmal zeigen. In der Testphase wurden die Probanden um eine einfache Kategorienverifikation gebeten ("Stimmt der gegebene Kategorienname mit dem abgebildeten Objekt überein oder nicht?"), während der ihre hirnelektrische Aktivität im Elektroenzephalogramm (EEG) aufgezeichnet wurde.

Die Analyse der Hirnaktivität zeigte, dass dieselben Objekte in Abhängigkeit von der sensu-motorischen Erfahrung beim Erwerb auf neuronaler Ebene unterschied-

lich verarbeitet werden: Ausschließlich in der Pantomimegruppe fand sich eine frühe Aktivierung über prämotorischen Arealen des Frontallappens, und dies bei einer Aufgabe (Kategorienverifikation), die keinen expliziten Rekurs auf handlungsrelevante Information erfordert. In Übereinstimmung mit den neurophysiologischen Daten schätzten die Probanden der Pantomime-Gruppe Handlungsinformation als bedeutsamer für die Konstituierung des Objektbegriffs ein als die Probanden der Zeige-Gruppe. Handlungsinformation trägt somit in Abhängigkeit von der spezifischen Lernerfahrung zur Begriffskonstituierung bei. Im Sinne der Modalitätsspezifität semantischer Systeme belegen diese Daten, dass sich die begrifflichen Repräsentationen durch die infolge der Lernerfahrung gebildeten Zellverbände in sensu-motorischen Hirnarealen konstituieren. Der Niederschlag in der subjektiven Einschätzung und auf neuronaler Ebene ist abhängig von dem Zugang, den man während des Erwerbs zu einem Objekt erhält. Diese Ergebnisse machen deutlich, dass die Etablierung handlungsrelevanter Begriffsmerkmale in Gehirn und Geist sinnvolle Handlungen mit den Objekten voraussetzt.

#### 2.1.2 Kontextabhängige Flexibilität von Begriffen

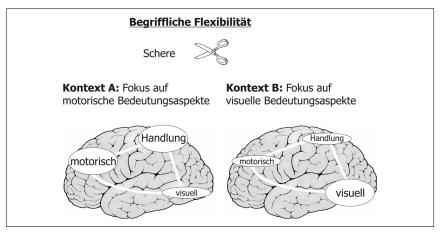


Abbildung 2: Flexibilität von Begriffen (nach Hoenig et al. 2008): Die Größe der ovalen Felder soll die differentiellen Beträge verteilter, jedoch anatomisch distinkter modalitätsspezifischer Hinnregionen unter zwei verschiedenen Kontextbedingungen repräsentieren. Die Annahme flexibler Begriffe geht davon aus, dass Begriffsmerkmale, die relevant sind für Motorik oder Bewegung, vorrangig rekrutiert werden, wenn der Kontext auf handlungsbezogene Bedeutungsaspekte fokussiert. Visuelle Begriffsmerkmale werden hingegen dann in relativ stärkerem Maße rekrutiert, wenn der kontextuelle Fokus auf die äußere Erscheinung der Objekte gerichtet ist. In jedem Fall wird angenommen, dass die entsprechenden Begriffsmerkmale in semantischen Feldern kodiert werden, welche sich in räumlich-anatomischer Nähe zu jenen Hirnarealen befinden, die an der Verarbeitung motorischer bzw. visueller Objektmerkmale beteiligt sind.

Im Gegensatz zur mittlerweile regen Debatte um die Modalitätsspezifität des semantischen Gedächtnisses blieb die zweite traditionelle Annahme, die die Stabilität begrifflichen Wissens postuliert, bislang weitgehend unhinterfragt. Ein wichtiger Grund hierfür dürfte sein, dass die Stabilitätsannahme nahezu nie explizit thematisiert wird, implizit aber immer schon in den grundlegendsten Prämissen bei der Untersuchung der neurokognitiven Architektur des begrifflichen Wissens enthalten ist.

Entgegen dieser Forderung wird bei der Annahme flexibler Begriffe gerade davon ausgegangen, dass der Kontext eines Begriffs von elementarer Bedeutung dafür ist, welche Begriffsmerkmale abgerufen werden (vgl. Abb. 2). Oder mit anderen Worten: Die subjektive bedeutungshaltige Rekonstruktion der Wirklichkeit hängt stark von der Situation ab. Auf der Ebene aktivierter Begriffsmerkmale sollte es demnach einen deutlichen Unterschied machen, ob es z. B. darum geht, was man mit einer Schere tun kann oder wie sie aussieht. In unserem Lebensalltag jedenfalls besitzt die Annahme flexibler Begriffe eine hohe intuitive Plausibilität (Barsalou et al. 2003, Kiefer 2005).

Zur Klärung der Frage, ob und wenn ja, in welchem Ausmaß modalitätsspezifische Hirnregionen zu einer flexiblen (kontextabhängigen) Verarbeitung begrifflichen Wissens beitragen, wurde jüngst eine Studie mittels fMRT und EEG durchgeführt, in der die beiden Faktoren Begriffskategorie (künstliche vs. natürliche Objekte) und semantische Verifikationsaufgabe (visuelle vs. handlungsrelevante Verifikation) gekreuzt wurden (Hoenig et al. 2008). Die Aufgabe bestand in einer Merkmalsverifikation zweier aufeinander folgender, visuell dargebotener Wörter (z. B. rot – Tomate oder schneiden – Messer vs. rund – Banane oder zupfen – Säge). Die Analyse der Bildgebungsdaten erbrachte signifikante Kreuzinteraktionen zwischen Aufgabe und Kategorie in visuellen und motorischen Hirnarealen. Zudem variierte das Interaktionsmuster in Abhängigkeit vom zugrunde liegenden modalitätsspezifischen Hirnareal (vgl. Abb. 3).

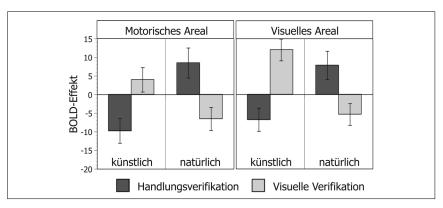


Abbildung 3: Dynamische Rekrutierung visueller und handlungsbezogener Begriffsmerkmale aus visuellen und motorischen Hirnarealen (nach Hoenig et al. 2008)

Wie die Ergebnisse zeigen, ist die Aktivierung in einem Areal besonders stark, wenn ein für eine Objektkategorie wenig relevantes Merkmal abgerufen werden soll, d. h. wenn der kognitive Aufwand hoch ist. Konkret bedeutet dies beispielsweise, dass ein motorisches Hirnareal am stärksten aktiviert wird, wenn ein handlungsrelevantes Merkmal (z. B. schneiden) verifiziert wird, und zwar von Objekten, für deren Begriffe Handlungsmerkmale nicht von vorrangiger Relevanz sind (z. B. Obiekte der natürlichen Kategorie Tier, da diese stärker durch ihr Aussehen definiert sind). Zusätzlich werden handlungsrelevante Begriffsmerkmale öfter aus der motorischen Hirnregion abgerufen, wenn Merkmale von Objekten verifiziert werden, deren Begriffe vorrangig durch Handlungsmerkmale (und damit durch motorisches Wissen) konstituiert werden. Dies ist etwa dann der Fall, wenn visuelle Merkmale (wie z. B. länglich) von Objekten der artifiziellen Kategorie Werkzeug verifiziert werden sollen. Diese Ergebnisse legen nahe, dass es sich bei Objektbegriffen um hochgradig flexible mentale Entitäten handelt, die sich unter anderem aus sensorischen und motorischen semantischen Merkmalen zusammensetzen. Diese Begriffsmerkmale werden in Abhängigkeit von ihrer Relevanz für die Objekte einer begrifflichen Kategorie einerseits und den kontextuellen Rahmenbedingungen andererseits dynamisch aus modalitätsspezifischen Hirnregionen rekrutiert.

#### 2.1.3 Der Klang der Begriffe im akustischen Kortex

Auf neurokognitiver Ebene lässt sich sogar zeigen, dass die Aktivität eines modalitätsspezifischen Hirnareals durch den subjektiven Erfahrungsgehalt der Probanden für ein bestimmtes Begriffsmerkmal moduliert wird (Kiefer et al. 2008): Probanden wurden im Kernspintomographen Wörter und sinnlose, aussprechbare Zeichenketten visuell dargeboten und bekamen die Aufgabe, eine simple lexikalische Entscheidung zu treffen. Sie mussten lediglich entscheiden, ob es sich bei den Zeichenketten um sinnvolle deutsche Substantive handelt oder nicht. Bei der Materialerstellung wurde sorgfältig darauf geachtet, dass sich die Wörter ausschließlich hinsichtlich ihrer Geräuschhaftigkeit voneinander unterscheiden, d. h. für die eine Hälfte sind Geräusche ein sehr relevantes Merkmal (z. B. Telefon, Säge, Hubschrauber), für die andere Hälfte jedoch haben sie keine oder nur eine geringe Relevanz (z. B. Lineal, Nagel, Hängematte). Beim Vergleich der Wörter in der funktionellen Hirnbildgebung fand man einen signifikanten Aktivitätsanstieg für geräuschbezogene Wörter in einem Areal der Hörrinde (begriffliche Aktivierung). Die begriffliche Aktivierung in dieser Hirnregion überlappt funktional-anatomisch mit einer ausgedehnteren Aktivierung, die beim Anhören realer Geräusche gefunden wurde (perzeptuelle Aktivierung).

Da für sämtliche Wörter von jedem Probanden individuelle Einschätzungen der Relevanz von Begriffsmerkmalen (akustisch, visuell und handlungsrelevant) vorlagen, wurde in einer dritten Analyse geprüft, ob die Aktivität in einem bestimmten Hirnareal systematisch moduliert wird durch die subjektive Einschätzung der Relevanz des akustischen Merkmals, nicht jedoch der anderen beiden Merkmale (vgl. Abb. 4). Eine für akustische Merkmale spezifische parametrische Modulation wurde lediglich in dem Bereich der Hörrinde gefunden, welcher sich in der funktional-anatomischen Überschneidungsregion aus der konzeptuellen und perzeptuellen Analyse befindet.

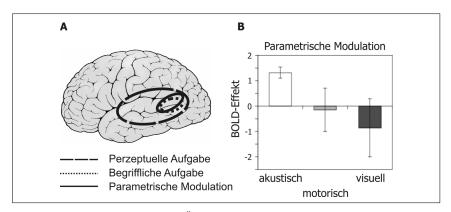


Abbildung 4: (A) Funktional-anatomische Überlappung von perzeptueller und begrifflicher Verarbeitung akustischer Information im Hörkortex; (B) Selektivität des Areals für die Kodierung akustischer Begriffsmerkmale (nach Kiefer et al. 2008)

Die erfahrungsabhängige Plastizität des Erwerbs begrifflichen Wissens, der neuronale Niederschlag subjektiver Begriffserfahrung und die kontextabhängige dynamische Rekrutierung begrifflicher Merkmale stellen ernste Herausforderungen für die Beibehaltung der Stabilitätsannahme dar. Sie machen sowohl sozial-konstruktive wie subjektive Aspekte bei der Gestaltung unserer gegenständlichen Welt deutlich. Gerade die Existenz dominanter Begriffsmerkmale von Objektkategorien (Handlungswissen für Artefakte und visuelles Wissen für natürliche Objekte) weist auf das Wirken intersubjektiver Konventionen hin, welche aber eine erfahrungsund kontextabhängige subjektive Ausformung erhalten. Begriffe sind die Bausteine, mit denen wir unsere Wirklichkeiten in sozialen Interaktionen unter den gegebenen gesellschaftlich-kulturellen Rahmenbedingungen verhandeln. Wir müssen schon – um mit Paul Feyerabend zu sprechen – wie Bildhauer der Realität auf die Welt einwirken, wenn wir neue (hier: mentale) Entitäten "entdecken" wollen. Der subjektseitige Erkenntnisanteil an der Wirklichkeitskonstruktion kann nicht einfach weggelassen werden, zumal er untrennbar mit ihren neuronalen Manifestationen verbunden ist. Die Studie zur erfahrungsabhängigen Plastizität des Begriffserwerbs macht aber auch deutlich, dass man sich der Welt in der richtigen Weise nähern muss. In manchen Fällen jedoch widersetzt sich das Material unseren Annäherungen (Feyerabend 1989).

### 3 Wege zu einer integrativen Psychologie

Es ist unstrittig, dass das Gehirn das biologische Substrat psychischer Phänomene ist. Die Erforschung von Gehirnprozessen liefert deshalb wichtige Information über deren Eigenschaften und die Architektur des kognitiven Systems. Neurowissenschaftliche Methoden ergänzen und erweitern das Methodenspektrum der Psychologie, indem sie physiologische Indikatoren dafür bereitstellen (Kiefer 2009).

Umgekehrt leistet neurowissenschaftlich orientierte psychologische Forschung einen wichtigen Beitrag zur Aufklärung der Funktionsweise des Gehirns. Unter diesem Aspekt kann eine mit der Biologie, der Medizin und der Neuroinformatik zusammenarbeitende Psychologie als Teildisziplin der Hirnforschung betrachtet werden (Born et al. 2003).

Zuweilen wird gefordert, dass sich die Psychologie vollständig der Hirnforschung widmen sollte, um so adäquate, naturwissenschaftliche Theorien für menschliches Erleben und Verhalten formulieren zu können (Churchland 1988, 1996). Psychologische Theorien, die sich auf das Gehirn als biologisches Substrat beziehen, hätten eine größere Erklärungskraft als solche, die allein mit theoretischen Konstrukten auf einer mentalen Ebene operieren.

Gleichzeitig wird zu Recht das Programm des radikalen Reduktionismus kritisiert, in dem psychologische Erklärungen ausnahmslos durch neurobiologische ersetzt werden (Fiedler et al. 2005, Kempf 2009a). Weil man gegenwärtig noch nicht verstehe, durch welche neuronalen Prozesse genau psychische Phänomene und Verhalten entstehen, könne noch nicht auf psychologische Begriffe verzichtet werden. Weiterhin erfassen die gegenwärtig vorhandenen und beim Menschen einsetzbaren Techniken zur Aufzeichnung der Hirnaktivität nur die kumulative Aktivität von großen Neuronenverbänden und erlauben lediglich ein sehr grobes Bild von den Vorgängen im Gehirn.

Selbst wenn die Hirnforschung in ferner Zukunft die neuronale Basis psychischer Phänomene besser erklären könnte als heute, wäre das Programm eines radikalen Reduktionismus wenig sinnvoll: Eine psychologische Erklärungsebene für das Denken, Erleben und Verhalten wäre auch in diesem Fall unverzichtbar, da die interessierenden Phänomene ohne einen Bezug auf psychologische Begriffe und Theorien weder beschreibbar noch erklärbar sind. Dies gilt vor allem dann, wenn deren subjektseitige, kulturelle und soziale Bedingungen betrachtet werden sollen. Aus den genannten Gründen ist das Programm eines radikalen neurobiologischen Reduktionismus zur Erklärung menschlichen Erlebens und Verhaltens letztendlich zum Scheitern verurteilt (Kempf 2009a, Kiefer 2009).

Eine verstärkte Vernetzung der natur- und kulturwissenschaftlichen Strömungen in der Psychologie ist deshalb nicht nur wünschenswert, sondern für viele Fragestellungen auch unabdingbar. Wenn man akzeptiert, dass die subjektive Welt des Menschen eine wichtige Bedingung für sein Handeln ist, dann kommt die Psychologie nicht ohne theoretische Begriffe aus, die mentale Vorgänge oder Zustände des Menschen bezeichnen. Soll zudem der Inhalt der subjektiven Welt nicht – wie es in der naturwissenschaftlichen Psychologie verbreitet ist – lediglich als Abbild formaler mentaler oder neuronaler Prozesse beschrieben, sondern vielmehr als Grund des Handelns rekonstruiert werden, dann ist der Einsatz kulturwissenschaftlicher Methoden unumgänglich, da diese darauf abzielen, die subjektive Sicht des Menschen auf seine Welt und ihre sozialen bzw. kulturellen Bedingungen zu erfassen. Umgekehrt liegen die Stärken der Methoden der naturwissenschaftlichen Psychologie in der feinen Analyse der Architektur menschlicher Kognition und der

dabei ablaufenden Prozesse. Eine integrative Psychologie beleuchtet deshalb im Idealfall ein und dasselbe psychische Phänomen mit natur- *und* kulturwissenschaftlichen Methoden und setzt die jeweiligen Befunde reziprok zueinander in Beziehung (Hönig 2009).

Wie könnte eine integrativ orientierte psychologische Forschung konkret aussehen? In Kapitel 2.1 haben wir einige Beispiele aus der Lern- und Gedächtnispsychologie aufgeführt, welche die Vision einer integrativen Psychologie bereits ansatzweise verwirklichen. Unsere Vorstellung von einer integrativen Psychologie wird besonders gut durch ein Projekt veranschaulicht, das wir zusammen mit Wilhelm Kempf durchführen, der seit vielen Jahren im Rahmen seiner friedensjournalistischen Forschung (Kempf et al. 2001) die Wirkung von unterschiedlichen Darstellungsperspektiven (Framings) auf die Rezeption von Zeitungstexten über aktuelle politische Konflikte (z. B. Golfkrieg, Krieg im ehemaligen Jugoslawien) untersucht. Die Nacherzählungen der Texte durch die Rezipienten werden dabei mittels qualitativer inhaltsanalytischer Verfahren analysiert, in deren Ergebnis bestimmt werden kann, wie die Rezipienten in Abhängigkeit von der Art des Framings den Konflikt subjektiv rekonstruieren (mentales Modell) (Kempf 2005, 2006a). Das Framing ist entweder eskalationsorientiert (unüberwindlicher Interessengegensatz; Gewalt als einziges adäguates Mittel der Konfliktlösung) oder deeskalationsorientiert (Interessengegensätze sind überbrückbar; Offenheit für Win-Win-Lösungen). Viele von Kempfs Untersuchungen zeigen, dass die Erinnerung an die berichteten Ereignisse, die Konfliktparteien und deren Konfliktverhalten von dieser Orientierung bestimmt wird (ebd.). Kempfs in der Tradition des kulturwissenschaftlichen Ansatzes stehende Studien geben detailliert Auskunft über den Inhalt der subjektiven Wirklichkeitsrekonstruktion, nicht aber darüber, welche kognitiven Prozesse durch das Framing beeinflusst werden. Werden dadurch bereits unmittelbar beim Lesen der Texte die Textverstehensprozesse moduliert oder erst Gedächtnisabrufprozesse bei der Nacherzählung? Aufschluss darüber können Experimente mit ereigniskorrelierten Potenzialen (EKP) geben, die uns Information über Textverstehensprozesse im Gehirn mit einer zeitlichen Auflösung von Millisekunden liefern, so dass Sprachverstehensprozesse fast in Echtzeit identifiziert werden können. Auch wenn eine solche EKP-Studie gegenwärtig noch in Planung ist (für eine vorbereitende Arbeit, vgl. Spohrs 2006), macht dieses Projekt doch bereits deutlich, wie kultur- und naturwissenschaftliche Ansätze theoriegeleitet in Abhängigkeit von der konkreten Fragestellung zu unterschiedlichen Zeiten des Forschungsprozesses eingesetzt werden können.

Kultur- und naturwissenschaftliche Ansätze in der Psychologie haben einen unterschiedlichen Schwerpunkt bei der Analyse psychischer Phänomene und ergänzen einander. Wir betrachten sie deshalb als komplementäre Forschungsparadigmen im Rahmen einer integrativen Psychologie und nicht als einander ausschließende. Seit ihrer Etablierung als akademische Disziplin ist die Psychologie mit dem Gegensatz zwischen der Subjektbezogenheit ihres Forschungsgegenstandes und der Notwendigkeit wissenschaftlicher Objektivität konfrontiert. Die Versuchung war und ist immer noch groß, das durch diesen Gegensatz aufgebaute Spannungsfeld

entweder durch die Aufgabe wissenschaftlicher Seriosität oder durch die Ausklammerung essenzieller, subjektbezogener Phänomenbereiche aufzulösen. Dieses Spannungsfeld aber zwischen Subjektivität und Objektivität, ja zwischen Kultur und Natur, ist der Psychologie jedoch inhärent und macht gerade die Faszination dieser Wissenschaft aus. Es lässt sich deshalb nicht per Dekret ohne Substanzverlust auflösen, sondern durch eine dialektische Synthese lediglich in seiner Brisanz entschärfen. Hierzu ist eine Überwindung der klassischen Grenzen zwischen Kultur- und Naturwissenschaft erforderlich (Kempf 2009a). Unsere hier skizzierte Vision einer integrativen Psychologie könnte einen gangbaren Weg für die Bewältigung dieser Herausforderung aufzeigen.

- Adorno, Theodor W. (1972). Einleitung. In: Adorno, Theodor W.; Albert, Hans & Dahrendorf, Ralf, Der Positivismusstreit in der deutschen Soziologie. Darmstadt: Luchterhand, S. 7-79.
- Adorno, Theodor W. (1957). Soziologie und empirische Forschung. In: GS 8. Frankfurt/M.: Suhrkamp, S. 196-216.
- Albert, Hans (1968, <sup>5</sup>1991). Traktat über kritische Vernunft. Tübingen: Mohr.
- American Psychiatric Association (1994). Diagnostic and statistical manual of mental disorders (4th ed.) (DSM IV). Washington, DC: American Psychiatric Press.
- Andersen, Erling & Madsen, Mette (1977). Estimating the Parameters of the Latent Population Distribution. Psychometrika, 42/3, pp. 357-374.
- Anderson, John R. (1983). The architecture of cognition. Hillsdayle, NJ: Lawrence Erlbaum Associates, Inc.
- Anderson, John R. & Bower, Gordon H. (1973). Human associative memory. Washington, DC: Hemisphere Press.
- Apel, Karl Otto (1979). Die Erklären-Verstehen-Kontroverse in transzendentalpragmatischer Sicht. Frankfurt/M.: Suhrkamp.
- Appelsmeyer, Heide & Billmann-Mahecha, Elfriede (Hrsg.) (2001). Kulturwissenschaft. Felder einer prozeßorientierten wissenschaftlichen Praxis. Weilerswist: Velbrück Wissenschaft.
- Aronson, Elliot; Wilson, Timothy & Akert, Robin (2008). Sozialpsychologie. München etc.: Pearson Studium.
- Axelrod, Robert (1988). Die Evolution der Kooperation. München: Oldenbourg.
- Baros, Wassilios (2001). Familien in der Migration. Frankfurt/M.: Peter Lang.
- Baros, Wassilios (2010). Innovative methodische Zugänge für qualitative Forschung im interkulturellen Kontext. In: Hagedorn, Jörg; Schurt, Verena; Steber, Corina & Waburg, Wiebke (Hrsg.), Ethnizität, Geschlecht, Familie und Schule. Heterogenität als erziehungswissenschaftliche Herausforderung. Wiesbaden: VS Verlag, S. 375-402.
- Baros, Wassilios (2012, i.E.). Globalisierungskritik bei jungen Erwachsenen als Medienrezipienten. Eine quantitative Inhaltsanalyse mittels Latent-Class-Analyse. In: conflict & communication online 11/2. URL: http://www.cco.regener-online.de
- Baros, Wassilios & Manafi, Georgia (2009). Migrantensein und Migrantenbewusstsein. Sozialpsychologische Rekonstruktion subjektiver Lebensorientierungen in prekären Lebenslagen. Diskussionsbeiträge des Forums Empirische Migrationsforschung. Unveröffentlichtes Manuskript.
- Baros, Wassilios & Reetz, Klaus-Dieter (2002). Sozialpsychologische Rekonstruktion und empirische Migrationsforschung. In: Conflict & communication online 1/2. URL: http://www.cco.regener-online.de

Barsalou, Larry W.; Simmons, W. Kyle; Barbey, Aron K. & Wilson, Christine D. (2003). Grounding conceptual knowledge in modality-specific systems. Trends in Cognitive Sciences, 7, pp. 84-91.

- Basisdaten Media Perspektiven (2006). URL (Zugriff 29.8.2007): http://www.media-perspektiven.de/1393.html
- Bateson, Gregory & Jackson, Don D. (1964). Some Varieties of Pathogenic Organization. In: Rioch, David McKenzie (Ed.), Disorders of Communication. Vol. 42, Research Publications. Association for Research in Nervous and Mental Disease, pp. 270-283.
- Baudrillard, Jean (1991). The Gulf War Did Not Take Place. Bloomington and Indianapolis: Indiana University Press.
- Baudrillard, Jean (2003). Der Geist des Terrorismus. Wien: Passagen Verlag.
- Baumert, Jürgen; Becker, Michael; Neumann, Marko & Nikolova, Roumiana (2009). Frühübergang in ein grundständiges Gymnasium – Übergang in ein privilegiertes Entwicklungsmilieu? Ein Vergleich von Regressionsanalyse und Propensity Score Matching. Zeitschrift für Erziehungswissenschaft, 12/2, S. 189-215.
- Beck, Aaron T. & Steer, Robert A. (1987). Manual for the Beck Depression Inventory. San Antonio, TX: The Psychological Corporation.
- Beck, Aaron T.; Steer, Robert A. & Brown, Gregory K. (1996). Beck Depression Inventory Second Edition. Manual. San Antonio, TX: The Psychological Corporation.
- Beck, Aaron T.; Steer, Robert A. & Garbin, Margery G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. Clinical Psychology Review, 8, pp. 77-100.
- Beck, Aaron T.; Ward, C.H.; Mendelson, M.; Mock, J. & Erbaugh, J. (1961). An inventory for measuring depression. Archives of General Psychiatry, 4, pp. 561-571.
- Beck, Ulrich (2007). Weltrisikogesellschaft. Frankfurt/M.: Suhrkamp.
- Becker, Jürgen (2002). Afghanistan: Der Krieg und die Medien. In: Albrecht, Ulrich (Hrsg.), Medien zwischen Krieg und Frieden. Baden-Baden: Nomos Verlagsgesellschaft, S. 142-172.
- Beckermann, Ansgar (1985). Handeln und Handlungserklärungen. In: Beckermann, Ansgar (Hrsg.), Analytische Handlungstheorie. Bd. 2, Frankfurt/M.: Suhrkamp, S. 7-84.
- Bennett, Lance & Paletz, David (Eds.) (1994). Taken by Storm. The Media, Public Opinion, and U.S. Foreign Policy in the Gulf War. Chicago: The University of Chicago Press.
- Benz, Carolyn R. & Newman, Isadore (1998). The Qualitative-Quantitative Research Methodology: Exploring the Interactive Continuum. Carbondale, IL: Southern Illinois University Press.
- Berger, Peter & Luckmann, Thomas (1969). Die gesellschaftliche Konstruktion der Wirklichkeit: Eine Theorie der Wissenssoziologie. Frankfurt/M.: Fischer.
- Biernatzki, E. William (2002). Terrorism and Mass Media. In: CSCC (Ed.), Communication research trends: a quarterly information service from the Centre for the Study of Communication and Culture London, 21/1. London: CSCC, pp. 1-19.
- Billmann-Mahecha, Elfriede (2001). Kann über die Methodenfrage die "Einheit" der Psychologie gerettet werden? Psychologie und Geschichte, 9 (3/4), S. 117-127.
- Billmann-Mahecha, Elfriede (2010). Auswertung von Zeichnungen. In: Mey, Günter & Mruck, Katja (Hrsg.), Handbuch Qualitative Forschung in der Psychologie. Wiesbaden: VS Verlag, S. 707-722.

Bläsi, Burkhard; Jaeger, Susanne; Kempf, Wilhelm & Möckel, Jutta (2004). A catalog of escalation- and de-escalation-oriented aspects of conflict coverage. Diskussionsbeiträge der Projektgruppe Friedensforschung Konstanz, Nr. 53.

- URL (letzter Zugriff 15.10.2010): http://www.uni-konstanz.de/FuF/SozWiss/fg-psy/agmeth/HTML-Kataloge/Checklist/index.htm
- Bock, R. Darrell & Aitkin, Murray (1981). Marginal Maximum Likelihood Estimation of Item Parameters: Application of an EM Algorithm. Psychometrika, 46/4, pp. 443-459.
- Bock, R. Darrell & Lieberman, Marcus (1970). Fitting a Response Model for n Dichotomously Scored Items. Psychometrika, 35/2, pp. 179-197.
- Boehnke, Klaus; Fuss, Daniel & Kindervater, Angela (2005). Peace psychology in Germany. Peace and Conflict: Journal of Peace Psychology, 11, pp. 229–237.
- Bönisch, Julia & Hauck, Mirjam (2009). Eurovision Song Contest 2009 in Moskau Swing-Pop mit Strip-Flop. Süddeutsche Zeitung vom 17.05.2009. URL (letzter Zugriff 15.10.2010): http://www.sueddeutsche.de/kultur/eurovision-songcontest-in-moskau-swing-pop-mit-strip-flop-1.450095
- Bonanno, A. George & Jost, T. John (2006). Conservative shift among highexposure survivors of the September 11th terrorist attacks. Basic and Applied Social Psychology, 28, pp. 311-323.
- Bonfadelli, Heinz (1993). Gewalt im Fernsehen. Gewalt durch Fernsehen. In: Bonfadelli, Heinz & Meier, Werner A. (Hrsg.), Krieg, Aids, Katastrophen. Gegenwartsprobleme als Herausforderung der Publizistikwissenschaft. Konstanz: UVK, S. 149-174.
- Born, Jan; Birbaumer, Niels; Elbert, Thomas; Güntürkün, Onur; Kirschbaum, Clemens & Rösler, Frank (2003). Biologische Psychologie 2010 Visionen zur Zukunft des Faches in der Psychologie. Psychologische Rundschau, 54, S. 120-124.
- Bortz, Jürgen (1984). Lehrbuch der empirischen Forschung für Sozialwissenschafter. Berlin: Springer.
- Bortz, Jürgen (2005). Statistik für Human- und Sozialwissenschaftler (6. Aufl.). Heidelberg: Springer.
- Bortz, Jürgen & Döring, Nicola (1995). Forschungsmethoden und Evaluation. 2., neu bearb. Auflage. Berlin: Springer.
- Bortz, Jürgen & Döring, Nicola (2002). Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. Berlin/Heidelberg: Springer.
- Bouman, Theo K. & Kok, A. R. (1987). Homogeneity of Beck's Depression Inventory (BDI): Applying Rasch analysis in conceptual exploration. Acta Psychiatrica Scandinavica, 76, pp. 568-573.
- Brandstätter, Eduard (1999). Confidence Intervals as an Alternative to Significance Testing. In: Methods of Psychological Research. Lengerich: Pabst Science Publishers. URL: http://www.mpr-online.de
- Brinkemper, V. Peter (2002). Die neue zivilsoldatische Mentalität. In: Palm, Goedart & Rötzer, Florian (Hrsg.), Medien. Terror. Krieg. Zum neuen Kriegsparadigma des 21. Jahrhunderts. Hannover: Heise, S. 207-222.
- Brown, Penelope & Levinson, Stephen C. (1987). Politeness: Some Universals in Language Usage. Cambridge: Cambridge University Press.
- Bücken, Hubert (Hrsg.) (2010). Unsere Lena Ein Mädchen verzaubert Europa. Düsseldorf: zeitgeistmedia.

Bühler, Karl (1965). Die Krise der Psychologie (Erstauflage: 1927). Stuttgart: Gustav Fischer.

- Bungard, Walter (1984). Psychologische Forschungsmethoden. In: Lück, Helmut E.; Müller, Rudolf & Rechtien, Wolfgang (Hrsg.), Geschichte der Psychologie. Ein Handbuch in Schlüsselbegriffen. München: Urban & Schwarzenberg, S. 40-49.
- Bungard, Walter; Schulz-Gambard, Jürgen & Antoni, Conny (1988). Zur Methodik der Angewandten Psychologie. In: Frey, Dieter; Graf Hoyos, Carl & Stahlberg, Dagmar (Hrsg.), Angewandte Psychologie. München: Psychologie Verlags Union, S. 588-606.
- Butler, John Scott (1985). The Statistical Bias of Numerically Integrated Statistical Procedures. Computers and Mathematics with Applications, 11/6, pp. 587-593.
- Cacioppo, John T. & Berntson, Gary G. (1992). Social psychological contributions to the decade of the brain. Doctrine of multilevel analysis. American Psychologist, 47, pp. 1019-1028.
- Canetti-Nisim, Daphna; Halperin, Eran; Sharvit, Keren & Hobfoll, Stevan (2009). A new stress-based model of political extremism: Personal exposure to terrorism, psychological distress, and exclusionist political attitudes. Journal of Conflict Resolution, 53, pp. 363-389.
- Campbell, Donald T. & Fiske, Donald W. (1959). Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix. Psychological Bulletin, 56/2, pp. 81-105.
- Carrier, Martin (2004). Interessen als Erkenntnisgrenzen? Die Wissenschaft unter Verwertungsdruck. In: Hogrebe, Wolfgang & Bromand, Joachim (Hrsg.), Grenzen und Grenz-überschreitungen. Berlin: Akademie-Verlag, S. 168-180.
- Carroll, J. Patrick; Wichman, L. Aaron & Arkin, A. Robert (2006). Security in the Aftermath of 9/11. Basic and Applied Social Psychology, 28/4, pp. 289–290.
- Christie, Daniel (2006). What is peace psychology the psychology of? Journal of Social Issues, 62, pp. 1-17.
- Christie, Daniel; Tint, Barbara; Wagner, Richard & Winter, Deborah DuNann (2008). Peace psychology for a peaceful world. American Psychologist, 63, pp. 540-552.
- Churchland, Patricia S. (1988). Reduction and the neurobiological basis of consciousness. In: Marcel, Anthony J. & Bisiach, Edoardo (Eds.), Consciousness in contemporary science. New York: Oxford University Press, pp. 273-304.
- Churchland, Patricia (1996). Die Neurobiologie des Bewußtseins. Was können wir von ihr lernen? In: Metzinger, Thomas (Hrsg.), Bewußtsein. Beiträge aus der Gegenwartsphilosophie (2. Aufl.). Paderborn: Schöningh, S. 463-490.
- Clark, David C.; Gibbons, Robert D.; Haviland, Mark G. & Hendryx, Michael S. (1993). Assessing the severity of depressive states in recently detoxified alcoholics. Journal of Studies on Alcohol, 54, pp. 107-114.
- Cohrs, J. Christopher & Boehnke, Klaus (2008). Social psychology and peace: An introductory overview. Social Psychology, 39, pp. 4-11.
- Cohrs, J. Christopher; Moschner, Barbara; Maes, Jürgen & Kielmann, Sven (2005a). Personal values and attitudes toward war. Peace and Conflict: Journal of Peace Psychology, 11/ 3, pp. 293-312.
- Cohrs, J. Christopher; Moschner, Barbara; Maes, Jürgen & Kielmann, Sven (2005b). The motivational bases of right-wing authoritarianism and social dominance orientation: Relations to values and attitudes in the aftermath of September 11, 2001. Personality and Social Psychology Bulletin, 31, pp. 1425-1434.

Coleman, Peter Thomas (2006). Conflict, complexity, and change: A meta-framework for addressing protracted, intractable conflicts – III. Peace and Conflict: Journal of Peace Psychology, 12, pp. 325-348.

- Coleman, Peter Thomas; Schneider, Adam; Adams, Douglass; Everett, Catherine James; Gameros, Timothy; Hammons, Lee; Orji, Cecil; Waugh, Ralph & Wicker, Richard (2005). Intragroup subgroup attitude clustering, external intervention, and intergroup interaction patterns: Toward a dynamical model of protracted intergroup conflict. Peace and Conflict Studies, 12, pp. 55-70.
- Crayen, Claudia; Geiser, Christian; Scheithauer, Herbert, & Eid, Michael (2011). Evaluating Interventions with Multimethod Data: A Structural Equation Modeling Approach. Structural Equation Modeling, 18, pp. 497-524.
- Crowson, H. Michael; Debacker, K. Teresa & Thoma, J. Stephen (2006). The Role of Authoritarianism, Perceived Threat, and Need for Closure or Structure in Predicting Post-9/11 Attitudes and Beliefs. The Journal of Social Psychology, 146/6, pp. 733-750.
- Dahrendorf, Ralf (1971). Zu einer Theorie des Sozialen Konflikts. In: Zapf, Wolfgang (Hrsg.), Theorien des sozialen Wandels. Köln: Kiepenheuer, S. 108-123.
- Danto, Arthur (1974). Analytische Psychologie der Geschichte. Frankfurt/M.: Suhrkamp.
- Davier, Matthias von (2001). WINMIRA 2001 user's guide. Kiel: IPN.
- Davies, G. Paul; Steele, M. Claude & Markus, R. Hazel (2008). A Nation Challenged: The Impact of Foreign Threat on America's Tolerance for Diversity. Journal of Personality and Social Psychology, 95/2, pp. 308-318.
- Dayan, Daniel & Katz, Elihu (1992). Media events: The live broadcasting of history. Cambridge, MA: Harvard University Press.
- Deci, Edward L. & Ryan, Richard M. (1985). Intrinsic motivation and self-determination in human behavior. New York: Plenum.
- Denzin, Norman K. (1970). The Research Act. Chicago: Aldine.
- Deutsch, Morton (1973). The Resolution of Conflict. Constructive and Destructive Processes. New Haven, CT: Yale University Press.
- Deutsch, Morton (1990). Sixty Years of Conflict. International Journal of Conflict Management 1, pp. 237-263.
- Dierstein, Jörg Michael (1995). Erklären oder Verstehen? Zur Konstruktion einer psychologischen Handlungstheorie. Münster New York: Waxmann.
- Dray, William (1985). Der Sinn der Handlungen. In: Beckermann, Ansgar (Hrsg.), Analytische Handlungstheorie. Bd. 2, Frankfurt/M.: Suhrkamp, S. 275-303.
- Egger, Andreas & Van Eimeren, Birgit (2008). Die Generation 60plus und die Medien. Zwischen traditionellen Nutzungsmustern und Teilhabe an der digitalen (R)evolution. Media Perspektiven, 11.
- Eichinger, Ulrike (2009). Zwischen Anpassung und Ausstieg. Perspektiven von Beschäftigten im Kontext der Neuordnung Sozialer Arbeit. Wiesbaden: VS Verlag.
- Eimer, Martin (1996). Kognitive Psychologie, Neurobiologie und das "Gehirn-Bewußtseins-Problem". In: Roth, Gerhard & Prinz, Wolfgang (Hrsg.), Kopf-Arbeit: Gehirnfunktionen und kognitive Leistungen. Heidelberg: Spektrum, Akademischer Verlag, S. 426-434.
- Elcheroth, Guy (2006). Individual-level and community-level effects of war trauma on social representations related to humanitarian law. European Journal of Social Psychology, 36, pp. 907-930.

Emler, Nicholas & Ohana, Jocelyne (1993). Studying Social Representations in Children: Just Old Wine in New Bottles? In: Breakwell, Glynis M. & Canter, David V. (Eds.), Empirical Approaches to Social Representations. Oxford: Clarendon Press, pp. 63-89.

- Entman, M. Robert (1993). Framing: Toward Clarification of a Fractured Paradigm. Journal of Communication, 43/4, pp. 51-58.
- Erb, Egon (1997). Gegenstands- und Problemkonstituierung: Subjektmodelle (in) der Psychologie. In: Groeben, Norbert (Hrsg.), Zur Programmatik einer sozialwissenschaftlichen Psychologie. Band I, Metatheoretische Perspektiven, 1. Halbbd.. Münster: Aschendorff, S. 139-239.
- Erzberger, Christian (1998). Zahlen und Wörter. Die Verbindung quantitativer und qualitativer Daten und Methoden im Forschungsprozeß. Weinheim: Deutscher Studien Verlag.
- Esses, M. Victoria; Dovidio, F. John & Hodson, Gordon (2002). Public Attitudes Toward Immigration in the United States and Canada in Response to the September 11, 2001 "Attack on America". The Society for the Study of Social Issues, 2/1.
- Fairclough, Norman (1995). Media Discourse. London: Routledge.
- Fairclough, Norman (2006). Language and Globalization. London: Routledge.
- Feyerabend, Paul K. (1989). Realism and the historicity of knowledge. Journal of Philosophy, 86, pp. 393-406.
- Fiedler, Klaus; Kliegl, Reinhold; Lindenberger, Ulman; Mausfeld, Rainer; Mummendey, Amélie, & Prinz, Wolfgang (2005). Psychologie im 21. Jahrhundert eine Standortbestimmung. Gehirn & Geist, 7-8, S. 56-60.
- Flick, Uwe (2010). Triangulation. In: Mey, Günter & Mruck, Katja (Hrsg.), Handbuch Qualitative Forschung in der Psychologie. Wiesbaden: VS Verlag, S. 278-289.
- Flick, Uwe; Kardorff, Ernst von; Keupp, Heiner; Rosensril, Lutz von & Wolff, Stephan (1991). Handbuch Qualitative Sozialforschung. Grundlagen, Konzepte, Methoden und Anwendungen. München Psychologie Verlags Union.
- Fodor, Jerry A. (2001). The mind doesn't work that way. Cambridge: MIT Press.
- Freire, Paulo (1981). Der Lehrer ist Politiker und Künstler. Neue Texte zur befreienden Bildungsarbeit. Reinbek: Rowohlt.
- Früh, Werner (1998). Inhaltsanalyse. Theorie und Praxis. Konstanz: UVK.
- Fukuyama, Francis (1989). The End of History? The National Interest, Summer.
- Galtung, Johan (2002). Peace Journalism A Challenge. In: Kempf, Wilhelm & Luostarinen, Heikki (Eds.), Journalism and the New World Order. Vol. II: Studying War and the Media. Göteborg: Nordicom.
- Gallese, Vittorio, & Lakoff, George (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. Cognitive Neuropsychology, 22, pp. 455-479.
- Gatherer, Derek (2006). Comparison of Eurovision Song Contest Simulation with Actual Results Reveals Shifting Patterns of Collusive Voting Alliances. Journal of Artificial Societies and Social Simulation, 9/2.
  - URL (letzter Zugriff 15.10.2010): http://jasss.soc.surrey.ac.uk/9/2/1.html
- Giddens, Anthony (1990). The Consequences of Modernity. Cambridge: Polity Press.
- Ginges, Jeremy & Atran, Scott (2008). Humiliation and the inertia effect: Implications for understanding violence and compromise in intractable intergroup conflicts. Journal of Cognition and Culture, 8, pp. 281–294.

Glaser, Barney G. & Strauss, Anselm L. (1967). The discovery of grounded theory. Strategies for qualitative research. Chicago: Aldine.

- Goffman, Erving (1959). The Presentation of Self in Everyday Life. Garden City, NY: Doubleday.
- Gollwitzer, Mario; Eid, Michael & Jürgensen, Ralph (2005). Response styles in the assessment of anger expression. Psychological Assessment, 17, pp. 56-69.
- Graumann, Carl F. (1973). Zur Lage der Psychologie. In: Reinert, Günter (Hrsg.), Bericht über den 27. Kongreß der Deutschen Gesellschaft für Psychologie in Kiel 1970. Göttingen: Hogrefe, S. 19-37.
- Graumann, Carl-Friedrich (1979). Die Scheu des Psychologen vor der Interaktion. Ein Schisma und seine Geschichte. Zeitschrift für Sozialpsychologie, 10, S. 284-304.
- Graumann, Carl-Friedrich (1985). Phänomenologische Analytik und experimentelle Methodik in der Psychologie das Problem der Vermittlung. In: Braun, Karl Heinz & Holzkamp, Klaus (Hrsg.), Subjektivität als Problem psychologischer Methodik. 3. Internationaler Kongreß Kritische Psychologie Marburg 1984. Frankfurt/M.: Campus, S. 38-59.
- Groeben, Norbert (1986). Die Herleitung von Erziehungszielen. In: Twellmann, Walter (Hrsg.), Handbuch Schule und Unterricht, Bd. 8.1. Düsseldorf: Schwann, S. 175-198.
- Groeben, Norbert (1999). Die metatheoretischen Merkmale einer sozialwissenschaftlichen Psychologie. In: Groeben, Norbert (Hrsg.), Zur Programmatik einer sozialwissenschaftlichen Psychologie. Band I: Metatheoretische Perspektiven, 2. Halbbd. Münster: Aschendorff, S. 311-404.
- Groeben, Norbert (2003). Problemaufriss einer Theoretischen Psychologie. In: Groeben, Norbert (Hrsg.), Zur Programmatik einer sozialwissenschaftlichen Psychologie. Band II: Objekttheoretische Perspektiven, 2. Halbbd. Münster: Aschendorff, S. 317-421.
- Groeben, Norbert (2006a). Die Überwindung der Grundlagen-Anwendungs-Implikation und Zielkriterien für praktische und epistemische Forschung eine wissenschaftstheoretische Perspektive. In: Brüggemann, Anne & Bromme, Rainer (Hrsg.), Entwicklung und Bewertung von anwendungsorientierter Grundlagenforschung in der Psychologie. Berlin: Akademie Verlag, S. 13-26.
- Groeben, Norbert (2006b). Gibt es Wege aus der selbstverschuldeten Irrelevanz des qualitativen off-streams? Forum Qualitative Sozialforschung, 7/4, Art. 34. URL: http://www.qualitative-research.net/fqs-texte/4-06/06-4-34-d.htm
- Groeben, Norbert & Scheele, Brigitte (1977). Argumente für eine Psychologie des reflexiven Subjekts. Paradigmawechsel vom behavioralen zum epistemologischen Menschenbild. Darmstadt: Steinkopff.
- Groeben, Norbert & Westmeyer, Hans (1981). Kriterien psychologischer Forschung. München: Juventa.
- Greve, Werner (1994). Handlungsklärung. Die psychologische Erklärung menschlicher Handlungen. Bern: Huber.
- Habermas, Jürgen (1965). Erkenntnis und Interesse. Merkur, 12, S. 1139-1153; auch in: Habermas, Jürgen (1969), Technik und Wissenschaft als "Ideologie". Frankfurt/M.: Suhrkamp, S. 146-168.
- Hafez, Kai (2007). The Myth of Media Globalization. Cambridge: Polity Press.
- Hartmann, Dirk & Werbik, Hans (2001). Über Reichweite und Grenzen einer naturwissenschaftlichen Psychologie. Handlung Kultur Interpretation, 10/1, S. 158-179.

Haußecker, Nicole (2007). Nachrichtenberichterstattung über Terrorismus. Eine Analyse der TV-Nachrichten über die Terroranschläge in Kenia. Conflict & communication online, 6/1. URL: http://www.cco.regener-online.de

- Hautzinger, Martin; Bailer, Maja, Worall, Hellgard & Keller, Ferdinand (1994). Beck-Depressions-Inventar (BDI). Bearbeitung der deutschen Ausgabe. Testhandbuch. Bern: Huber.
- Hautzinger, Martin, Keller, Ferdinand & Kühner, Christine (2006). BDI-II. Beck-Depressions-Inventar Revision Manual. Frankfurt: Harcourt Test Services.
- Hebb, Donald O. (1949). The organization of behavior. A neuropsychological theory. New York: Wiley.
- Heinze, Thomas & Klusemann, Hans-W. (1979). Ein biographisches Interview als Zugang zu einer Bildungsgeschichte. In: Baacke, Dieter & Schultze, Theodor (Hrsg.), Aus Geschichten lernen. München: Juventa, S. 195-201.
- Held, Josef (1989). Subjektbezogene Jugendforschung aus kritisch-psychologischer Sicht. In: Held, Josef (Hrsg.), Subjektbezogene Jugendforschung. Argument Sonderband 174. Hamburg: Argument, S. 16-41.
- Herrmann, Theo (1985). Allgemeine Sprachpsychologie: Grundlagen und Probleme (2. Aufl., Weinheim: Beltz, Psychologie-Verlags-Union, 1994 ed.). München: Urban & Schwarzenberg.
- Herrmann, Theo, & Grabowski, Joachim (1994). Sprechen: Psychologie der Sprachproduktion. Heidelberg: Spektum, Akademischer Verlag.
- Higgins, Kathleen M. (2009). Kitsch. In: Davies, Stephen; Higgins, Kathleen M.; Hopkins, Robert, Stecker, Robert & Cooper, David E. (Eds.), A Companion to Aesthetics, 2nd Edition, Wiley-Blackwell, pp. 393-396.
- Hönig, Klaus (2009). Konvergenz von kulturwissenschaftlichen und naturwissenschaftlichen Paradigmen der Psychologie. In: Kempf, Wilhelm & Kiefer, Markus (Hrsg.), Forschungsmethoden der Psychologie. Zwischen naturwissenschaftlichem Experiment und sozialwissenschaftlicher Hermeneutik. Bd. 3: Natur und Kultur. Berlin: regener, S. 271-323.
- Hoenig, Klaus; Sim, Eun-Jin; Bochev, Viktor; Herrnberger, Bärbel & Kiefer, Markus (2008). Conceptual flexibility in the human brain: Dynamic recruitment of semantic maps from visual, motion and motor-related areas. Journal of Cognitive Neuroscience, 20, pp. 1799-1814.
- Hoffman, Bruce (2002). Terrorismus der unerklärte Krieg. Bonn: Bundeszentrale für Politische Bildung.
- Hoffmann, Joachim (1986). Die Welt der Begriffe. Psychologische Untersuchungen zur Organisation des menschlichen Wissens. Weinheim: Beltz, Psychologie-Verlags-Union.
- Holzkamp, Klaus (1970). Zum Problem der Relevanz psychologischer Forschung für die Praxis. Psychologische Rundschau, 21, S. 1-22 (bei Markard zit. nach dem Nachdruck in Holzkamp, Klaus (2009). Schriften V (hrsg. Von Haug, Frigga et al), Hamburg: Argument, S. 15-40.
- Holzkamp, Klaus (1972a). Kritische Psychologie. Frankfurt: Fischer.
- Holzkamp, Klaus (1972b). Verborgene anthropologische Voraussetzungen der allgemeinen Psychologie. In: Holzkamp, Klaus, Kritische Psychologie. Vorbereitende Arbeiten. Frankfurt/M.: Fischer, S. 35-73. Nachdruck in: Holzkamp, Klaus (2009). Schriften V (hrsg. von Haug, Frigga et al.). Hamburg: Argument: S. 41-82.

Holzkamp, Klaus (1981). Nachwort. In: Klaus Holzkamp, Theorie und Experiment in der Psychologie (2. Aufl., S. 275-278). Berlin: de Gruyter.

- Holzkamp, Klaus (1983; 2. Aufl. 1985). Grundlegung der Psychologie. Frankfurt/M.: Campus.
- Holzkamp, Klaus (1986). Die Verkennung von Handlungsbegründungen als empirische Zusammenhangsannahmen in sozialpsychologischen Theorien: Methodologische Fehlorientierung infolge von Begriffsverwirrung. Zeitschrift für Sozialpsychologie, 17, S. 216-238; zit. nach dem Nachdruck in: Forum Kritische Psychologie, 19, 1987, S. 23-58.
- Holzkamp, Klaus (1991). Was heißt "Psychologie vom Subjektstandpunkt"? Forum Kritische Psychologie, 28, S. 5-19.
- Holzkamp, Klaus (1993). Lernen. Subjektwissenschaftliche Grundlegung. Frankfurt/M.: Campus.
- Holzkamp, Klaus (1996, posthum). Psychologie: Selbstverständigung über Handlungsbegründungen alltäglicher Lebensführung. Forum Kritische Psychologie, 36, S. 7-74.
- Hornsey, Matthew & Hogg, Michael (2000). Assimilation and diversity: An integrative model of subgroup relations. Personality and Social Psychology Review, 4, pp. 143-156.
- Huber, Oswald (1987). Das psychologische Experiment: Eine Einführung. Bern: Huber.
- Huck, Lorenz (2009). Jugendliche "Intensivtäter/innen" in Berlin Kriminelle Karrieren und Präventionsmöglichkeiten aus der Sicht der betroffenen Subjekte. Hamburg: Argument.
- Huddy, Leonie; Feldman, Stanley & Weber, Christopher (2007). The political consequences of perceived threat and felt insecurity. The Annals of the American Academy, 614, pp. 131-153.
- Humphreys, Glyn W. & Forde, Emer M. E. (2001). Hierarchies, similarity, and interactivity in object recognition: "Category-specific" neuropsychological deficits. Behavioral and Brain Sciences, 24, pp. 453-509.
- Huntington, Samuel (1996). The Clash of Civilizations and the Remaking of World. New York: Simon & Schuster.
- Hussy, Walter; Schreier, Margit & Echterhoff, Gerald (2010). Forschungsmethoden in Psychologie und Sozialwissenschaften für Bachelor. Berlin: Springer.
- Irtel, Hans (1987). A conjoint Grassmann structure for testing the additivity of binocular color mixtures. Journal of Mathematical Psychology, 31, pp. 192-202.
- Irtel, Hans (1987). On specific objectivity as a concept in measurement. In: Roskam, Edward E. & Suck, Reinhard (Eds.), Progress in mathematical Psychology. Vol. 1. Amsterdam: North-Holland, pp. 35-45.
- Irtel, Hans (1987). Intensity invariance in binocular brightness processing. In: Roskam, Edward E. & Suck, Reinhard (Eds.), Progress in mathematical Psychology. Vol. 1. Amsterdam: North-Holland, pp. 325-335.
- Irtel, Hans & Schmalhofer, Franz (1982). Psychodiagnostik auf Ordinalskalenniveau: Meßtheoretische Grundlagen, Modelltest und Parameterschätzung. Archiv für Psychologie, 134, S. 197-218.
- Jaeger, Susanne (2009). Nachrichtenmedien als Ressource für Frieden und Versöhnung. Inhaltsanalytische Pressestudien zur westdeutschen Berichterstattung über Frankreich nach dem Zweiten Weltkrieg. Berlin: regener.
- Jahoda, Marie; Lazarsfeld, Paul & Zeisl, Hans (1975). Die Arbeitslosen von Marienthal. Ein soziographischer Versuch über die Wirkungen langandauernder Arbeitslosigkeit (Erstauflage: 1933). Frankfurt/M.: Suhrkamp.

- James, William (1890). Principles of Psychology. New York: Dove Publishers.
- Janich, Peter; Kambartel, Friedrich & Mittelstraß, Jürgen (1974). Wissenschaftstheorie als Wissenschaftskritik. Frankfurt/M.: Aspekte Verlag.
- Jasinskaja-Lahti, Inga; Liebkind, Karmela & Perhoniemi, Riku (2006). Perceived discrimination and well-being: A victim study of different immigrant groups. Journal of Community and Applied Social Psychology, 16, pp. 267-284.
- Jenkins, Brian (1975). International Terrorism. Los Angeles: Crescent Publication.
- Joe, Harry (1990). Multivariate concordance. Journal of Multivariate Analysis, 35, pp. 12-30.
- Joe, Harry (1997). Multivariate Models and Dependence Concepts. London: Chapman & Hall.
- Jones, Edward (1998). Major developments in five decades of social psychology. In: Gilbert, Daniel; Fiske, Susan & Lindzey, Gardner (Eds.), The handbook of social psychology (Vol. 1, 4th ed.). New York: Oxford University Press, pp. 3-57.
- Jüttemann, Gerd (Hrsg.) (1985). Qualitative Forschung in der Psychologie. Grundfragen, Verfahrensweisen, Anwendungsfelder. Weinheim: Beltz.
- Jüttemann, Gert (1991). Die Bedeutung regulativer Motivationen für die Analyse des Subjekts. In: Jüttemann, Gert (Hrsg.), Individuelle und soziale Regeln des Handelns. Beiträge zur Weiterentwicklung geisteswissenschaftlicher Ansätze in der Psychologie. Heidelberg: Asanger, S. 93-103.
- Kaldor, Mary (1999). New and old wars: organized violence in a global era. Cambridge: Polity Press.
- Kamlah, Wilhelm & Lorenzen, Paul (1967). Logische Propädeutik. Mannheim: Bibliographisches Institut.
- Keller, Ferdinand (1998). Probabilistische Testmodelle zur Abschätzung der Homogenität von BDI-Subskalen. In: Stieglitz, Rolf-Dieter; Fähndrich, Erdmann & Möller, Hans-Jürgen (Hrsg.), Syndromale Diagnostik psychischer Störungen. Göttingen: Hogrefe, S. 129-136.
- Keller, Ferdinand (2000). Methoden zur Unterscheidung dimensionaler und kategorialer latenter Strukturen in der Depressionsdiagnostik. In: Reinecke, Jost & Tarnai, Christian (Hrsg.), Angewandte Klassifikationsanalyse in den Sozialwissenschaften. Münster: Waxmann, S. 178-207.
- Keller, Ferdinand; Hautzinger, Martin & Kühner, Christine (2008). Zur faktoriellen Struktur des deutschsprachigen BDI-II. Zeitschrift für Klinische Psychologie und Psychotherapie, 37, S. 245-254.
- Keller, Ferdinand; Hoffmann, Markus & Weithmann, Gerd (2002). Psychometrische Eigenschaften des Beck-Depressionsinventars (BDI) bei Alkoholabhängigen. Zeitschrift für Klinische Psychologie und Psychotherapie, 31, S. 169-177.
- Keller, Ferdinand & Kempf, Wilhelm (1993). Stabilität und Validität latenter Klassen im Beck-Depressionsinventar (BDI). Psychiatrische Praxis, 20 (Suppl.), S. 68-71.
- Keller, Ferdinand & Kempf, Wilhelm (1997). Some latent trait and latent class analyses of the Beck-Depression-Inventory (BDI). In: Rost, Jürgen & Langeheine, Rolf (Eds.), Applications of latent trait and latent class models in the social sciences. Münster: Waxmann, pp. 314-323.
- Kellner, Douglas (1992). The Persian Gulf TV War. Boulder Col.: Westview Press.
- Kempf, Wilhelm (1978). Konfliktlösung und Aggression. Bern/Stuttgart: Hans Huber.

Kempf, Wilhelm (1991). Transkulturelle Verständnisbildung als Methodenproblem. In: Kempf, Wilhelm (Hrsg.), Verdeckte Gewalt. Hamburg: Argument Verlag, S. 119-132.

- Kempf, Wilhelm (1992a). Zum Verhältnis von qualitativen und quantitativen Methoden in der psychologischen Forschung. Forum Kritische Psychologie, 29, S. 89-108.
- Kempf, Wilhelm (1992b). Ein pragmatischer Ansatz in der statistischen Theorie psychologischer Testscores. Zeitschrift für Differentielle und Diagnostische Psychologie, 4, S. 269-275.
- Kempf, Wilhelm (1993). Konflikteskalation durch autonome Prozesse. In: Kempf, Wilhelm; Frindte, Wolfgang; Sommer, Gert & Spreiter, Michael (Hrsg.), Gewaltfreie Konfliktlösungen. Heidelberg: Asanger, S. 53-69.
- Kempf, Wilhelm (Hrsg.) (1994a). Manipulierte Wirklichkeiten. Medienpsychologische Untersuchungen der bundesdeutschen Presseberichterstattung im Golfkrieg. Münster: Lit.
- Kempf, Wilhelm (1994b). Zum Empirebezug subjektwissenschaftlicher Erklärungen. Forum Kritische Psychologie, 34, S. 54-60.
- Kempf, Wilhelm (1994c). Diskussionsbeitrag. In: Brandtstädter, Jochen et al., Zur Problematik des Empiriebezugs psychologischer Theorien. Forum Kritische Psychologie, 34, S. 5-79.
- Kempf, Wilhelm (1996). Konfliktberichterstattung zwischen Eskalation und Deeskalation. Ein sozialpsychologisches Modell. Wissenschaft und Frieden, 14, S. 51-54.
- Kempf, Wilhelm (2001). News media and conflict escalation a comparative study of the Gulf War coverage in American and European media. In: Nohrstedt, Stig Arne & Ottosen, Rune (Eds.), Journalism and the New World Order. Vol. I: Gulf War, National News Discourses and Globalization. Göteborg: Nordicom.
- Kempf, Wilhelm (2002). Integration of Quantitative and Qualitative Content Analysis in Media Research. In: Kempf, Wilhelm & Luostarinen, Heikki (Eds.), Journalism and the New World Order. Vol. II. Studying War and the Media. Göteborg: Nordicom, pp. 145-172.
- Kempf, Wilhelm (2003a). Forschungsmethoden der Psychologie. Zwischen naturwissenschaftlichem Experiment und sozialwissenschaftlicher Hermeneutik. Bd. I: Theorie und Empirie. Berlin: regener.
- Kempf, Wilhelm (2003b). Konstruktive Konfliktberichterstattung Ein sozialpsychologisches Forschungs- und Entwicklungsprogramm. Conflict & communication online, 2/2. URL: http://www.cco.regener-online.de
- Kempf, Wilhelm (2005). Two experiments focusing on de-escalation oriented coverage of post-war conflicts. Conflict & communication online, 4/2. URL: http://www.cco.regener-online.de
- Kempf, Wilhelm (2006a). The impact of text framing on the understanding of political conflict. Diskussionsbeiträge der Projektgruppe Friedensforschung Konstanz, Nr. 59. URL: http://www.regener-online.de
- Kempf, Wilhelm (2006b). Social constructionism and its implications for critical media studies. Conflict & communication online, 4/2. URL: http://www.cco.regener-online.de
- Kempf, Wilhelm (2007). Peace journalism: A tightrope between advocacy journalism and constructive conflict coverage. Conflict & communication online, 6/2. URL: http://www.cco.regener-online.de
- Kempf, Wilhelm (2008). Forschungsmethoden der Psychologie. Zwischen naturwissenschaftlichem Experiment und sozialwissenschaftlicher Hermeneutik. Bd. II: Quantität und Qualität. Berlin: regener.

Kempf, Wilhelm (2009a). Psychologie als Natur- und Kulturwissenschaft. In: Kempf, Wilhelm & Kiefer, Markus (Hrsg.), Forschungsmethoden der Psychologie. Zwischen naturwissenschaftlichem Experiment und sozialwissenschaftlicher Hermeneutik. Bd. III: Natur und Kultur. Berlin: regener, S. 17-48.

- Kempf, Wilhelm (2009b). Die soziale Konstruktion der Wirklichkeit In: Kempf, Wilhelm & Kiefer, Markus (Hrsg.), Forschungsmethoden der Psychologie. Zwischen naturwissenschaftlichem Experiment und sozialwissenschaftlicher Hermeneutik. Band III: Natur und Kultur. Berlin: regener, S. 51-99.
- Kempf, Wilhelm (2010). Patterns of criticizing Israel and their relation to modern Anti-Semitism. Conflict & communication online, 9/1. URL: http://www.cco.regener-online.de
- Kempf, Wilhelm; Baros, Wassilios & Regener, Irena (2000). Sozialpsychologische Rekonstruktion Integration quantitativer und qualitativer Forschungsmethoden in der psychologischen Konflikt- und Friedensforschung. Forum Qualitative Sozialforschung, 1,2, Art. 12. URL: http://nbn-resolving.de/urn:nbn:de:0114-fqs0002128
- Kempf, Wilhelm & Kiefer, Markus (Hrsg.) (2009a). Forschungsmethoden der Psychologie. Zwischen naturwissenschaftlichem Experiment und sozialwissenschaftlicher Hermeneutik. Band III: Natur und Kultur. Berlin: regener.
- Kempf, Wilhelm & Kiefer, Markus (2009b). Einleitung. In: Kempf, Wilhelm & Kiefer, Markus (Hrsg.), Forschungsmethoden der Psychologie. Zwischen naturwissenschaftlichem Experiment und sozialwissenschaftlicher Hermeneutik. Bd. III: Natur und Kultur. Berlin: regener, S. 7-13.
- Kempf, Wilhelm & Luostarinen, Heikki (Eds.) (2002), Journalism and the New World Order. Vol. II: Studying War and the Media. Göteborg: Nordicom.
- Kempf, Wilhelm & Reimann, Michael (1994). Informationsbedürfnis und Mediengebrauch während des Golfkrieges. In: Kempf, Wilhelm (Hrsg.), Manipulierte Wirklichkeiten. Medienpsychologische Untersuchungen der bundesdeutschen Presseberichterstattung im Golfkrieg. Münster: LIT, S. 47-57.
- Kempf, Wilhelm; Reimann, Michael & Luostarinen, Heikki (1996). Qualitative Inhaltsanalyse von Kriegspropaganda und Kritischem Friedensjournalismus. Diskussionsbeiträge der Projektgruppe Friedensforschung, Projekt 13/85, Universität Konstanz Nr. 32/1996.
- Kempf, Wilhelm; Reimann, Michael & Luostarinen, Heikki (2001). New world order rhetoric in US and European media. In: Nohrstedt, Stig Arne & Ottosen, Rune (Eds.), Journalism and the New World Order. Vol. I: Gulf War, National News Discourses and Globalization. Göteborg: Nordicom, pp. 125-148.
- Kempf, Wilhelm; Reimann, Michael & Luostarinen Heikki (2002). The presentation of alternative ways to settle the Gulf Conflict in German, Norwegian and Finnish media. In: Kempf, Wilhelm & Luostarinen, Heikki (Eds.), Journalism and the New World Order. Vol. II: Studying War and the Media. Göteborg: Nordicom.
- Keppler, Angela (1994). Tischgespräche: Über Formen kommunikativer Vergemeinschaftung am Beispiel der Konversation in Familien. Frankfurt/M.: Suhrkamp.
- Keppler, Angela (2000). Verschränkte Gegenwarten. Medien- und Kommunikationssoziologie als Untersuchung kultureller Transformationen. In: Münch, Richard; Jauß, Claudia & Stark, Carsten (Hrsg.), Soziologische Revue Sonderheft 5 (Soziologie 2000). München: Oldenbourg, S. 140-152.
- Kerlinger, Fred N. (1975). Grundlagen der Sozialwissenschaft (2 Bde.). Weinheim: Beltz.
- Kiefer, Markus (1999). Die Organisation des semantischen Gedächtnisses. Bern: Huber.

Kiefer, Markus (2005). Repetition priming modulates category-related effects on event-related potentials: Further evidence for multiple cortical semantic systems. Journal of Cognitive Neuroscience, 17, pp. 199-211.

- Kiefer, Markus (2008). Zusammenwirken kognitiver Systeme: Kognitionspsychologische und neurophysiologische Befunde zur Rolle des semantischen Gedächtnisses bei der Informationsverarbeitung. Psychologische Rundschau, 59, S. 87-97.
- Kiefer, Markus (2009). Neurowissenschaftliche Methoden. In: Kempf, Wilhelm & Kiefer, Markus (Hrsg.), Forschungsmethoden der Psychologie. Zwischen naturwissenschaftlichem Experiment und sozialwissenschaftlicher Hermeneutik. Bd. III: Natur und Kultur. Berlin: regener, S. 207-267.
- Kiefer, Markus; Sim, Eun-Jin; Herrnberger, Bärbel; Grothe, Jo & Hoenig, Klaus (2008). The sound of concepts: Four markers for a link between auditory and conceptual brain systems. The Journal of Neuroscience, 28, pp. 12224-12230.
- Kiefer, Markus; Sim, Eun-Jin; Liebich, Sarah; Hauk, Olaf & Tanaka, James W. (2007). Experience-dependent plasticity of conceptual representations in human sensory-motor areas. Journal of Cognitive Neuroscience, 19, pp. 525-542.
- Kiefer, Markus & Spitzer, Manfred (2001). The limits of a distributed account of conceptual knowledge. Trends in Cognitive Sciences, 5, pp. 469-471.
- Kimmel, Allan J. (1996). Ethical issues in behavioral research. Oxford: Blackwell Publishers.
- Klein, Tovah; Devoe, R. Ellen; Miranda-Julian, Claudia & Linas, Keri (2009). Young Children's Respones To September 11th: The New York City Experience. Infant Mental Health Journal, 30/1, pp. 1-22.
- Klüver, Jürgen (1979). Kommunikative Validierung. In: Heinze, Theodor (Hrsg.), Lebensweltanalyse von Fernstudenten. Theoretische und methodologische Überlegungen zum Typus hermeneutisch-lebensgeschichtlicher Forschung. Fernuniversität Hagen. Hagen, S. 69.
- Knappertsbusch, Felix & Kelle, Udo (2010). "Mutterland des nomadisierendes Finanzkapitals"
  Zum Verhältnis von Antiamerikanismus und Antisemitismus vor dem Hintergrund der Finanzkrise. In: Heitmeyer, Wilhelm (Hrsg.), Deutsche Zustände. Folge 8. Frankfurt/M.: Suhrkamp, S. 144-163.
- Kochinka, Alexander & Werbik, Hans (1997). Logische Propädeutik und Wissenschaftstheorie. In: Straub, Jürgen; Kempf, Wilhelm & Werbik, Hans (Hrsg.), Psychologie. Eine Einführung. München: dtv, S. 42-67.
- Kölbl, Carlos & Billmann-Mahecha, Elfriede (2005). Die Gruppendiskussion. Schattendasein einer Methode und Plädoyer für ihre Entdeckung in der Entwicklungspsychologie. In: Mey, Günter (Hrsg.), Handbuch Qualitative Entwicklungspsychologie. Köln: Kölner Studienverlag, S. 321-350.
- Köller, Olaf; Baumert, Jürgen & Rost, Jürgen (1998). Zielorientierungen: Ihr typologischer Charakter und ihre Entwicklung im frühen Jugendalter. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 30/3, S.128-138.
- König, Eckard (1975). Theorie der Erziehungswissenschaft, Bd. 2. Normen und ihre Rechtfertigung. München: Fink.
- Kosslyn, Stephan M. (2005). Vorwort. In: Libet, Benjamin (Hrsg), Mind Time. Wie das Gehirn Bewusstsein produziert. Frankfurt/M.: Suhrkamp, S. 9-16.
- Kriesberg, Louis (1973). The Sociology of Social Conflicts. Englewood Cliffs, NJ: Prentice-Hall.

Kriz, Jürgen; Lück, Helmut E. & Heidbrink, Horst (1987). Wissenschaftstheorie und Erkenntnistheorie. Eine Einführung für Psychologen und Humanwissenschaftler. Opladen: Leske + Budrich.

- Kubinger, Klaus D. & Draxler, Clemens (2007). Probleme bei der Testkonstruktion nach dem Rasch-Modell. Diagnostica, 53, S. 131-143.
- Kühner, Christine; Bürger, Christine; Keller, Ferdinand & Hautzinger, Martin (2007). Reliabilität und Validität des revidierten Beck Depressionsinventars (BDI-II) Befunde aus deutschsprachigen Stichproben. Der Nervenarzt, 78, S. 651-656.
- Kuhn, Thomas S. (1967). Die Struktur wissenschaftlicher Revolutionen. Frankfurt/M.: Suhrkamp.
- Kulka, Tomas (1996). Kitsch and Art. University Park, Pennsylvania: Pennsylvania State University Press.
- Kundera, Milan (1987). Die unerträgliche Leichtigkeit des Seins. Frankfurt/M.: Fischer Taschenbuch.
- Kutschera, Franz von (1972). Wissenschaftstheorie I und II. München: Fink.
- Lakoff, Georg, & Johnson, Mark (1999). Philosophy in the Flesh: The embodied mind and its challenge to Western thought. New York: Basic Books.
- Lamnek, Siegfried (1989). Qualitative Sozialforschung. Bd. 2. Methoden und Techniken. München: Psychologie Verlags Union.
- Lamnek, Siegfried (1995). Qualitative Sozialforschung. Band 1: Methodologie. 3., korr. Auflage. Weinheim: Psychologie Verlags Union.
- Lazarsfeld, Paul F. & Henry, Neil W. (1968). Latent structure analysis. Boston: Houghton Mifflin Co.
- Le Guern, Philippe. (2002). From National Pride to Global Kitsch: the Eurovision Song Contest. Lille: University of Lille.
  - URL (letzter Zugriff 15.10.2010): http://wjfms.ncl.ac.uk/leguWJ.htm
- Liang, Kung-Yee & Zeger, Scott L. (1986). Longitudinal Data Analysis Using Generalized Linear Models. Biometrika, 73/1, pp. 13-22.
- Libet, Benjamin (2005). Mind Time. Wie das Gehirn Bewusstsein produziert. Frankfurt/M.: Suhrkamp.
- Lienert, Gustav A. & Raatz, Ulrich (1994). Testaufbau und Testanalyse. 5. Aufl. Weinheim:
- Lorenzen, Paul & Schwemmer, Oswald (1975). Konstruktive Logik, Ethik und Wissenschaftstheorie, 2. Aufl. Mannheim: Bibliographisches Institut.
- Lück, Helmut E.; Grünwald, Harald; Geuter, Ulfried & Rechtien, Wolfgang (1987). Sozialge-schichte der Psychologie. Eine Einführung. Opladen: Leske + Budrich.
- Luhmann, Niklas (1984). Soziale Systeme: Grundriß einer allgemeinen Theorie. Frankfurt/M.: Suhrkamp.
- Luhmann, Niklas (1995). Die Kunst der Gesellschaft. Frankfurt/M.: Suhrkamp.
- Luria, Alexandr R. (1966a). Higher cortical functions in man. London: Tavistock Publications.
- Luria, Alexandr R. (1966b). Human Brain and Psychological Processes. New York: Harper and Row.
- Luria, Alexandr R. (1968). The mind of a mnemonist: A little book about a vast memory. Cambridge, Massachusetts: Harvard University Press.

Luria, Alexandr R. (1970). Traumatic aphasia. Its syndromes, psychology and treatment. Den Haag und Paris: Walter de Gruyter Inc.

- Lurija, Alexandr R. (1991). Der Mann, dessen Welt in Scherben ging: zwei neurologische Geschichten. Reinbek bei Hamburg: Rowohlt.
- Lutz, Rainer (2001). Das Marburger Untersuchungsinstrumentarium (MUM). Unveröffentlichtes Manuskript. Fachbereich Psychologie der Universität Marburg.
- Maier, Michaela (2003). Analysen deutscher Fernsehnachrichten 1992-2001. In: Ruhrmann, Georg; Woelke, Jens; Maier, Michaela & Diehlmann, Nicole (Hrsg.), Der Wert von Nachrichten im deutschen Fernsehen. Opladen: Leske + Budrich, S. 61-98.
- Maier, Michaela & Stengel, Karin (2007). Bedeutung des Faktors Visualität für die Nachrichtenberichterstattung über internationale Krisen. In: Glaab, Sonja (Hrsg.), Medien und Terrorismus Auf den Spuren einer symbiotischen Beziehung. Berlin: Berliner Wissenschaftsverlag, S. 127-138.
- Maiers, Wolfgang (1994). Subjektiv begründetes Handeln als psychologische Analyseeinheit. In: Hoefert, Hans-Wolfgang & Klotter, Christoph (Hrsg.), Neue Wege der Psychologie. Eine Wissenschaft in der Veränderung. Heidelberg: Asanger, S. 57-79.
- Mair, Patrick & Hatzinger, Reinhold (2007). Extended Rasch modeling: The eRm package for the application of IRT models in R. Journal of Statistical Software, 20, pp. 1-20.
- Markard, Morus (1984). Einstellung. Kritik eines sozialpsychologischen Grundkonzepts. Frankfurt/M.: Campus.
- Markard, Morus (2000). Verbale Daten, Entwicklungsfigur, Begründungsmuster, Theorien-prüfung: Methodische Probleme und Entwicklungen in der Projektarbeit. In: Markard, Morus & Ausbildungsprojekt Subjektwissenschaftliche Berufspraxis, Kritische Psychologie und studentische Praxisforschung. Wider Mainstream und Psychoboom. Konzepte und Erfahrungen des Ausbildungsprojekts Subjektwissenschaftliche Berufspraxis an der FU Berlin. Hamburg: Argument, S. 227-250.
- Markard, Morus (2009, 3. Aufl. 2010). Einführung in die Kritische Psychologie. Hamburg: Argument.
- Markard, Morus (2010). Kritische Psychologie: Forschung vom Standpunkt des Subjekts. In: Mey, Günter & Katja Mruck (Hrsg.), Handbuch Qualitative Forschung in der Psychologie. Wiesbaden: VS Verlag, S. 166-181.
- Martín-Baró, Ignacio (1994). Writings for a liberation psychology. Cambridge, MA: Harvard University Press.
- Mattes, Peter (1985). Die Psychologiekritik der Studentenbewegung. In: Ash, Mitchell G. (Hrsg.), Geschichte der deutschen Psychologie im 20. Jahrhundert. Opladen: Westdeutscher Verlag, S. 286-313.
- Matthes, Jörg (2007). Framing-Effekte. Zum Einfluss der Politikberichterstattung auf die Einstellungen der Rezipienten. München: Reinhard Fischer.
- Mayring, Philipp (1996). Einführung in die Qualitative Sozialforschung. Eine Anleitung zu qualitativem Denken. 3. Auflage. Weinheim: Psychologie Verlags Union.
- Mayring, Philipp (1999). Qualitativ orientierte Forschungsmethoden in der Unterrichtswissenschaft: Ein Anwendungsbeispiel aus der Lernstrategieforschung. In: Unterrichtswissenschaft, S. 292-309.
- Mayring, Philipp (2001). Kombination und Integration qualitativer und quantitativer Analyse. Forum Qualitative Sozialforschung, 2/1.
  - URL: http://www.qualitative-research.net/index.php/fgs/article/view/967

McCullagh, Peter (1983). Quasi-Likelihood Functions. The Annals of Statistics, 11/1, pp. 59-67.

- McGrath, Joseph E. & Brinberg, David (1984). Alternative paths for research: Another view of the basic versus applied distinction. Applied Social Psychology Annual, 5, pp. 109-129.
- McLuhan, Marshall (1964). Understanding Media: The Extensions of Man. New York: Mc-Graw-Hill.
- McLuhan, Marshall & Fiore, Quentin (1984). Das Medium ist Massage. Frankfurt, Berlin, Wien: Ullstein.
- Mertens, Wolfgang & Fuchs, Gudrun (1978). Krise der Sozialpsychologie? Zur Krisendiskussion über die theoretischen und methodischen Grundlagen der Sozialpsychologie. München: Ehrenwirth.
- Messmer, Heinz (2003). Der soziale Konflikt. Kommunikative Emergenz und systemische Reproduktion. Stuttgart: Lucius & Lucius.
- Métraux, Alexander (1985). Der Methodenstreit und die Amerikanisierung der Psychologie in der Bundesrepublik 1950-1970. In: Ash, Mitchell G. (Hrsg.), Geschichte der deutschen Psychologie im 20. Jahrhundert. Opladen: Westdeutscher Verlag, S. 225-251.
- Mey, Günter & Mruck, Katja (Hrsg.) (2010). Handbuch Qualitative Forschung in der Psychologie. Wiesbaden: VS Verlag.
- Miebach, Bernhard (2009). Prozesstheorie. Analyse, Organisation und System. Wiesbaden: VS Verlag.
- Moser, Heinz (1995). Grundlagen der Praxisforschung. Freiburg im Breisgau: Lambertus.
- Mouffe, Chantal (2008). Das demokratische Paradox. Wien: Turia & Kant.
- Muchow, Martha & Muchow, Hans Heinrich (1998). Der Lebensraum des Großstadtkindes (Erstauflage: 1935). Weinheim: Juventa.
- Müller-Freienfels, Richard (1929). Die Hauptrichtungen der Psychologie. Leipzig: Quelle & Meyer.
- Musallam, Naira; Coleman, Peter & Nowak, Andrzej (2010). Understanding the spread of malignant conflict: A dynamical systems perspective. Peace and Conflict: Journal of Peace Psychology, 16, pp. 127-151.
- Neuman, W. Russel; Just, Marion & Crigler, N. Ann (1992). Common knowledge: news and the construction of political meaning. Chicago: University of Chicago Press.
- Nicholls, John G. (1978). The Development of the Concepts of Effort and Ability, Perception of Academic Attainment, and the Understanding That Difficult Tasks Require More Ability. Child development, 49, pp. 800-814.
- Nohrstedt, Stig Arne (Ed.) (2011). Communicating Risks Towards the Threat Society? Göteborg: Nordicom.
- Nohrstedt, Stig Arne; Höijer, Birgitta & Ottosen, Rune (2002). Kosovokonflikten, medierna och medlidandet, Stockholm: Styrelsen för Psykologiskt Försvar Rapport 190.
- Nohrstedt, Stig Arne & Ottosen, Rune (2001). Journalism and the New World Order. Vol. I: Gulf War, National News Discourse and Globalization. Göteborg: Nordicom.
- Nohrstedt, Stig Arne & Ottosen, Rune (Eds.) (2004). U.S. and the Others. Global Media Images on the "War on Terror". Göteborg: Nordicom.
- Nohrstedt, Stig Arne & Ottosen, Rune (Eds.) (2005). Global War Local News. Media Images of the Iraq War. Göteborg: Nordicom.

Nohrstedt, Stig Arne & Ottosen, Rune (2010). A New Information Order – But What Order? In: Weibull, Lennart (Ed.), The Nordic Countries and the World Perspectives from research and Media Communication. A book for Ulla Carlsson. Göteborg: University of Gothenburg.

- Nohrstedt, Stig Arne & Tveiten, Odger (2002). News, Discourse, Rhetoric, Propaganda. Conflict Journalism from a Multi-Methodological Perspective. In: Kempf, Wilhelm & Luostarinen, Heikki (Eds.), Journalism and the New World Order. Vol. II: Studying War and the Media. Göteborg: Nordicom.
- Oevermann, Ulrich (1974). Die falsche Kritik an der kompensatorischen Erziehung. In: Neue Sammlung 14, S. 537-568.
- Opp, Karl-Dieter (1976). Methodologie der Sozialwissenschaften: Einführung in die Probleme ihrer Theorienbildung. (2. Aufl.). Reinbek: Rowohlt.
- Ottosen, Rune & Figenschou, Ustad, T. (2007). September 11 in Norwegian Media: Images of the local threat. In: Pludowski, Thomasz (Ed.), How the World's News Media reacted to 9/11. Essays from around the globe. Spokane: Marquette books.
- Pajala, Mari (2007). Finlande: zero points? Der Eurovision Song Contest in den finnischen Medien. Köln: Saxa Verlag.
- Paluck, Elizabeth Levy (2009). Reducing intergroup prejudice and conflict using the media: A field experiment in Rwanda. Journal of Personality and Social Psychology, 96, pp. 574-587.
- Pannenberg, Markus & Spiess, Martin (2009). GEE Estimation of a Two-Equation Panel Data Model with an Application to Wage Dynamics and the Incidence of Profit-Sharing in West Germany. AStA Advances in Statistical Analysis, 93, pp. 427-447.
- Parsons, Talcott (1968). Interaction: Social Interaction. In: Sills, David L. (Ed.), International Encyclopedia of the Social Sciences, Vol. 7. New York: Macmillan & Free Press, pp. 429-440.
- Popper, Karl Raimund (1966). Logik der Forschung. [Wien 1935/1998.] 2. Aufl. Tübingen:
- Prinz, Wolfgang (1996). Bewußtsein und Ich-Konstitution. In: Roth, Gerhard & Prinz, Wolfgang (Hrsg.), Kopf-Arbeit: Gehirnfunktionen und kognitive Leistungen. Heidelberg: Spektrum, S. 451-467.
- Pruitt, Dean & Rubin, Jeffrey (1986). Social Conflict, Escalation, Stalemate and Settlement. New York: Random House.
- Pulvermüller, Friedemann (2005). Brain mechanisms linking language and action. Nature Reviews: Neuroscience, 6, pp. 576-582.
- Quilty, Lena C., Zhang, K.Anne & Bagby, R.Michael (2010). The latent symptom structure of the Beck Depression Inventory-II in outpatients with major depression. Psychological Assessment, 22, pp. 603-608.
- Rapoport, Anatol (1974). Konflikte in der von Menschen gemachten Umwelt. Darmstadt: Schwarz.
- Rasch, Georg (1960). Probabilistic Models for Some Intelligence and Attainment Tests. Copenhagen: The Danish Institute of Educational Research. (Erweiterte Auflage, 1980. Chicago: The University of Chicago Press.)
- Rasch, Georg (1961). On General Laws and the Meaning of Measurement in Psychology. Proceedings of the VI. Berkeley Symposium on Mathematical Statistics and Probability, Vol. IV. Berkeley: University of California Press, pp. 321-333.

Raykoff, Ivan (2007). Camping on the Borders of Europe. In: Raykoff, Ivan & Tobin, Robert D. (Eds.), A song for Europe: popular music and politics in the Eurovision Song Contest. Aldershot: Ashgate Publishing Limited, pp. 1-12.

- Reetz, Klaus-Dieter (2007). Migration und schulischer Misserfolg italienischer Kinder. Berlin: regener.
- Reinders, Heinz. 2006. Kausalanalysen in der Längsschnittforschung. Das Crossed-Lagged-Panel-Design. Diskurs Kindheits- und Jugendforschung, 1/4, S. 569-587.
- Reinecke, Jost (2005). Strukturgleichungsmodelle in den Sozialwissenschaften. München: Oldenbourg.
- Richter, Paul; Werner, Joachim & Bastine, Reiner (1994). Psychometrische Eigenschaften des Beck-Depressionsinventars (BDI): Ein Überblick. Zeitschrift für Klinische Psychologie, 23, S. 3-19.
- Richter, Paul; Werner, Joachim; Heerlein, Andres; Kraus, Alfred & Sauer, Heinrich (1998).

  On the validity of the Beck Depression Inventory: A review. Psychopathology, 31, pp. 160-168.
- Robertson, Roland (1992). Globalization. Social Theory and Global Culture. London: Sage.
- Rohracher, Hubert (1963). Einführung in die Psychologie. 8. Aufl. Wien: Urban & Schwarzenberg.
- Rost, Jürgen (1995). Die testdiagnostische Erfassung von Typen. In: Pawlik, K. (Hrsg.), Bericht über den 39. Kongreß der DGPs in Hamburg. Göttingen: Hogrefe, S. 392-398.
- Rost, Jürgen (1997). Intelligenz. In: Kempf, Wilhelm; Straub, Jürgen & Werbik, Hans: Psychologie. Eine Einführung. Grundlagen, Methoden, Forschungsfelder. München: DTV.
- Rost, Jürgen (1998). Drei Thesen zum Konzept qualitativer Forschungsmethoden. Zeitschrift für Didaktik der Naturwissenschaften, 4/3, S. 35-42.
- Rost, Jürgen (1999). Was ist aus dem Rasch-Modell geworden? Psychologische Rundschau, 50, S. 140-156.
- Rost, Jürgen (2002). Zeitgeist und Moden empirischer Analysemethoden. Zuma Nachrichten, 8, S. 21-31.
- Rost, Jürgen (2004). Lehrbuch Testtheorie, Testkonstruktion (2. Aufl.). Bern: Huber.
- Rost, Jürgen; Carstensen, Claus & Davier, Matthias von (1999). Sind die Big Five Rasch-skalierbar? Diagnostica, 45, S. 119-127.
- Santor, Darcy A.; Ramsay, Jim O. & Zuroff, David C. (1994). Nonparametric item analyses of the Beck Depression Inventory: Evaluating gender item bias and response option weights. Psychological Assessment, 6, pp. 255-270.
- Sarris, Viktor (1990). Methodologische Grundlagen der Experimentalpsychologie. 1: Erkenntnisgewinnung und Methodik. München/Basel: PVU.
- Schäffer, Burkhard (2006). Gruppendiskussion. In: Ayaß, Ruth & Bergmann, Jörg (Hrsg.), Qualitative Methoden der Medienforschung. Reinbeck bei Hamburg: Rowohlt, S. 115-145.
- Scheiblechner, Hartmann (1995). Isotonic ordinal probabilistic models (ISOP). Psychometrika, 60, 281-304.
- Scheiblechner, Hartmann (1999). Additive conjoint isotonic probabilistic Models (ADISOP). Psychometrika, 64, pp. 295-316.

Scheiblechner, Hartmann (2003). Nonparametric IRT: testing the bi-isotonicity of isotonic probabilistic models (ISOP). Psychometrika, 68, pp. 79-96.

- Scheiblechner, Hartmann (2007). A unified nonparametric IRT model for d-dimensional psychological test data (d-ISOP). Psychometrika, 72, pp. 43-67.
- Scheiblechner, Hartmann & Lutz, Rainer (2009). Die Konstruktion eines optimalen eindimensionalen Tests mittels nichtparametrischer Testtheorie (NIRT) am Beispiel des MR SOC. Diagnostica, 55/1, Göttingen: Hogrefe, S. 41-54.
- Scheufele, Bertram (2003). Frames Framing Framing-Effekte. Theoretische und methodische Grundlegung sowie empirische Befunde zur Nachrichtenproduktion. Wiesbaden: Westdeutscher Verlag.
- Schilling, Stephen & Bock, R. Darrell (2005). High-Dimensional Maximum Marginal Likelihood Item Factor Analysis by Adaptive Quadrature. Psychometrika, 70/3, pp. 1-23.
- Schmid, P. Alex & de Graaf, Jenny (1982). Violence as Communication. Insurgent Terrorism and the western News Media. London, Beverly Hills: Sage.
- Schneckener, Ulrich (2006). Transnationaler Terrorismus. Frankfurt/M.: Suhrkamp.
- Schreier, Margrit (2006a). Experimentelle/quasi-experimentelle Untersuchungsplanung. In: Groeben, Norbert & Hurrelmann, Bettina (Hrsg.), Empirische Unterrichtsforschung in der Literatur- und Lesedidaktik. Weinheim: Juventa, S. 307-342.
- Schreier, Margrit (2006b). Qualitatives Untersuchungsdesign. In: Groeben, Norbert & Hurrelmann, Bettina (Hrsg.), Empirische Unterrichtsforschung in der Literatur- und Lesedidaktik. Weinheim: Juventa, S. 343-359.
- Schreier, Margrit (2006c). Qualitative Verfahren der Datenerhebung. In: Groeben, Norbert & Hurrelmann, Bettina (Hrsg.), Empirische Unterrichtsforschung in der Literatur- und Lesedidaktik. Weinheim: Juventa, S. 399-420.
- Schreier, Margit & Odağ, Özen (2010). Mixed Methods. In: Mey, Günter & Mruck, Katja (Hrsg.), Handbuch Qualitative Forschung in der Psychologie. Wiesbaden: VS Verlag, S. 263-277.
- Schuler, Heinz. (1980). Ethische Probleme psychologischer Forschung. Göttingen: Hogrefe.
- Schurz, Gerhard (Hrsg.) (1990). Erklären und Verstehen in der Wissenschaft. München: Oldenbourg.
- Schweiger, Wolfgang & Brosius, Hans-Bernd (2003). Eurovision Song Contest beeinflussen Nachrichtenfaktoren die Punktvergabe durch das Publikum? M&K 51, 2/2003, S. 271-294.
- Schwemmer, Oswald (1976). Theorie der rationalen Erklärung. München: Beck.
- Skitka, J. Linda; Bauman, W. Christopher; Aramovich, P. Nicholas & Morgan, G. Scott (2006). Confrontational and Preventative Policy Responses to Terrorism: Anger Wants a Fight and Fear Want "Them" to Go Away. Basic and Applied Social Psychology, 28/4, pp. 375– 384.
- Sommer, Gerd & Fuchs, Albert (Hrsg.) (2004). Krieg und Frieden. Handbuch der Konfliktund Friedenspsychologie. Weinheim: Beltz PVU.
- Spiess, Martin (2006). Estimation of a Two-Equation Panel Model With Mixed Continuous and Ordered Categorical Outcomes and Missing Data. Journal of the Royal Statistical Society: Series C (Applied Statistics), 55/4, pp. 525-538.
- Spiess, Martin (1995). Parameterschätzung in Regressionsmodellen mit abhängigen diskreten endogenen Variablen. Konstanz: Hartung-Gorre.

Spiess, Martin & Hamerle, Alfred (2000). A Comparison of Different Methods for the Estimation of Regression Models with Correlated Binary Responses. Computational Statistics & Data Analysis, 32, pp. 439-455.

- Spiess, Martin & Hamerle, Alfred (1996). On the properties of GEE estimators in the presence of invariant covariates. Biometrical Journal, 38, pp. 931-940.
- Spohrs, Monika (2006). Über den Nachrichtenwert von Friedensjournalismus Ergebnis einer experimentellen Studie. Conflict & communication online, 5/1.

  URL: http://www.cco.regener-online.de
- Staub, Ervin (2007). Preventing violence and terrorism and promoting positive relations between Dutch and Muslim communities in Amsterdam. Peace and Conflict: Journal of Peace Psychology, 13/3, pp. 333-360.
- Staub, Ervin & Vollhardt, Johanna (2008). Altruism born of suffering: The roots of caring and helping after experiences of personal and political victimization. American Journal of Orthopsychiatry, 78, pp. 267-280.
- Straub, Jürgen (1999). Handlung, Interpretation, Kritik. Berlin: de Gruyter.
- Strauss, Anselm & Corbin, Juliet (1996). Grounded Theory. Grundlagen Qualitativer Sozial-forschung. Weinheim: Psychologie Verlags Union (engl. Orig. 1990).
- Steer, Robert A.; Ball, Roberta; Ranieri, William F. & Beck, Aaron T. (1999). Dimensions of the Beck Depression Inventory-II in clinically depressed outpatients. Journal of Clinical Psychology, 55, pp. 117-128.
- Tausch, Nicole; Hewstone, Miles & Roy, Ravneeta (2009). The relationships between contact, status and prejudice: An integrated threat theory analysis of Hindu-Muslim relations in India. Journal of Community & Applied Social Psychology, 19, pp. 83-94.
- Teddlie, Charles & Tashakkori, Abbas (2009). Foundations of Mixed Methods Research. Los Angeles: Sage.
- Thiel, Christian (1980). Argumentation. In: Mittelstraß, Jürgen (Hrsg.), Enzyklopädie Philosophie und Wissenschaftstheorie, Bd. 1. Mannheim: J. B. Metzler.
- Thomae, Hans & Feger, Hubert (1969). Hauptströmungen der neueren Psychologie. Frankfurt/M.: Akademische Verlags-Gesellschaft.
- Thomae, Hans (1977). Psychologie in der modernen Gesellschaft. Hamburg: Hoffmann und Campe.
- Tiedemann, Joachim & Faber, Günter (1995). Mädchen im Mathematikunterricht: Selbstkonzept und Kausalattributionen im Grundschulalter. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 27/1, S. 61-71.
- Tiedemann, Joachim (2000). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. Journal of Educational Psychology, 92, pp. 144-151.
- Trope, Yaakov & Liberman, Nira (2010). Construal-level theory of psychological distance. Psychological Review, 117, pp. 440-463.
- Tsfati, Yariv & Weimann, Gabriel (2002). www.terrorism.com: Terror on the Internet. Studies in Conflict & Terrorism, 25, pp. 317-332.
- Tyler, Lorraine K. & Moss, Helen E. (2001). Towards a distributed account of conceptual knowledge. Trends in Cognitive Sciences, 5, pp. 244-252.

Van Zelst, Willeke; De Beurs, Edwin & Smit, H. Johannes (2003). Effects of the September 11th on symptoms of PTSD on community-dwelling older persons in the Netherlands. International Journal of Geriatric Psychiatry, 18/2, p. 190.

- Vallacher, Robin; Coleman, Peter; Novak, Andrzej & Bui-Wrzosinska, Lan. (2010). Rethinking intractable conflict. The perspective of dynamical systems. American Psychologist, 65, pp. 262-278.
- Velmans, Max (1991). Is human information processing conscious? Behavioral and Brain Sciences, 14, pp. 651-726.
- Vollhardt, Johanna & Bilali, Rezarta (2008). Social psychology's contribution to the psychological study of peace: A review. Social Psychology, 39, pp. 12-25.
- Vollhardt, Johanna Ray (2009a). Altruism born of suffering and prosocial behavior following adverse life events: A review and conceptualization. Social Justice Research, 22, pp. 53-97.
- Vollhardt, Johanna Ray (2009b). The role of victim beliefs in the Israeli-Palestinian conflict: Risk or potential for peace? Peace and Conflict: Journal of Peace Psychology, 15, pp.135-159.
- Vollhardt, Johanna Ray (2011). "Crime against humanity" or "Crime against Jews"? The importance of acknowledgment in construals of the Holocaust for intergroup relations (eingereicht, Journal of Social Issues).
- Von Eye, Alexander (1994). Zum Verhältnis zwischen qualitativen und quantitativen Methoden in der empirisch-pädagogischen Forschung. In: Olechowski, Richard & Rollett, Brigitte (Hrsg.), Aspekte empirisch-pädagogischer Forschung quantitative und qualitative Methoden. Bericht über die 49. Tagung der Arbeitsgruppe für empirisch-pädagogische Forschung. Frankfurt/M.: Lang, S. 24-45.
- Vuchinich, Samuel (1987). Starting and Stopping Spontaneous Family Conflicts. Journal of Marriage and the Family, 49, pp. 591-601.
- Vygotsky, Lev S. (1927) The Historical Meaning of the Crisis in Psychology: A Methodological Investigation. In Rieber, Robert W. & Wollock, Jeffrey (1997), The collected works of Lev Vygotsky. Vol. 3, Chapter 15, Springer US, pp. 233-345.
- Waldmann, Peter (2005). Terrorismus. Provokation der Macht. Hamburg: Murmann.
- Watzlawick, Paul; Beavin, Janet H. & Jackson, Don D. (1982). Menschliche Kommunikation: Formen, Störungen, Paradoxien. 6. unveränd. Aufl. Bern: Huber.
- Weimann, Gabriel & Brosius, Hans-Bernd (1989). Attraktivität von Gewalt. Über welche internationalen Terroranschläge berichten die Medien? Publizistik, 34, S. 329.
- Weiner, Bernard (1986). An Attributional Theory of Motivation and Emotion. New York: Springer.
- Werbik, Hans (1985). "Psychonomie" und "Psychologie". Zur Notwendigkeit der Unterscheidung zweier Wissenschaften. In: Burrichter, Clemens; Inhetveen, Rüdiger & Kötter, Rudolf (Hrsg.), Technische Rationalität und rationale Heuristik. Paderborn: Schöningh, S. 109-121.
- Werbik, Hans & Appelsmeyer, Heide (1999). Handlungsbegriff und Perspektivität. Eine Diskussion traditioneller Perspektiven psychologischer Erkenntnisgewinnung. In: Straub, Jürgen & Werbik, Hans (Hrsg.), Handlungstheorie. Begriff und Erklärung des Handelns im interdisziplinären Diskurs. Frankfurt/M.: Campus, S. 74-92.

Werthes, Sascha; Kim, Richard & Conrad, Christoph (2002). Die Terrorkrise als Medienereignis? In: Schicha, Christian & Brosda, Carsten (Hrsg.), Medien und Terrorismus. Reaktionen auf den 11. September 2001. Münster, Hamburg: Lit, S. 80-93.

- Westermann, Rainer (2000). Wissenschaftstheorie und Experimentalmethodik. Göttingen: Hogrefe.
- Wohlrapp, Harald (2009). Der Begriff des Arguments. Über die Beziehungen zwischen Wissen, Forschen, Glaube, Subjektivität und Vernunft. 2. Aufl. Würzburg: Königshausen & Neumann.
- Wolther, Irving (2006). "Kampf der Kulturen". Der Eurovision Song Contest als Mittel national-kultureller Repräsentation. Würzburg: Königshausen & Neumann.
- World Health Organization (1992). The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines. Geneve: World Health Organization.
- Wright, Georg Henrik von (1974). Erklären und Verstehen. Frankfurt/M.: Fischer Athenäum.
- Wundt, Wilhelm (1874). Grundzüge der physiologischen Psychologie. Leipzig: Engelmann.
- Wundt, Wilhelm (1911). Vorlesungen über die Menschen- und Thierseele (Erstauflage: 1863). Hamburg und Leipzig: Leopold Voss.
- Wundt, Wilhelm (1912). Elemente der Völkerpsychologie. Grundlinien einer psychologischen Entwicklungsgeschichte der Menschheit. Leipzig: Kröner.

Baros, Wassilios, Prof. Dr., geb. 1969. Ass. Professor für Interkulturelle Bildung an der Democritus Universität of Thrace/GR. Z. Z. Vertretungsprofessor (Erziehung und Migration) an der Goethe-Universität Frankfurt und Gastprofessor (Lehrstuhl für Pädagogik) an der Universität Augsburg. Arbeitsschwerpunkte: Methodologie und Methoden interkultureller Bildungsforschung, political literacy und Migration, Theorie und Empirie transformatorischer Bildungsprozesse.

Kontakt: Goethe-Universität, Fachbereich Erziehungswissenschaften, Institut für Allgemeine Erziehungswissenschaft, Fach 111, Senckenberganlage 15, 60054 Frankfurt/M. eMail: vbaros@eled.duth.gr

Billmann-Mahecha, Elfriede, Prof. Dr., geb. 1951. Professorin für Psychologie an der Leibniz Universität Hannover. Arbeitsschwerpunkte: Entwicklungspsychologie, Kulturpsychologie, Methoden.

Kontakt: Institut für Pädagogische Psychologie, Leibniz Universität Hannover Schloßwender Str. 1, 30159 Hannover eMail: billmann@psychologie.uni-hannover.de URL: http://www.psychologie.uni-hannover.de/elfriede billmann.html

Frindte, Wolfgang, Prof. Dr. phil. habil., geb. 1951. Professor für Kommunikationspsychologie an der Friedrich-Schiller-Universität Jena.

Kontakt: Friedrich-Schiller-Universität Jena, Institut für Kommunikationswissenschaft Abteilung Kommunikationspsychologie, Ernst-Abbe-Platz 8, 07743 Jena eMail: wolfgang.frindte@uni-jena.de

Groeben, Norbert, Prof. Dr. phil. habil. (Psychologie), habil. (Allgemeine Literaturwissenschaft), Dipl.Psych., M.A. (Neuere Dt. Philologie), geb. 1944. Univ.-Prof. a. D. für Allgemeine und Kulturpsychologie, Universität Köln; Hon.-Prof. für Allgemeine und empirische Literaturwissenschaft, Universität Mannheim. Arbeitsschwerpunkte: Sprach- und Denkpsychologie, psychologische Anthropologie, Wissenschafts- und Literaturtheorie, empirische Literaturwissenschaft und Medienpsychologie.

Kontakt: Peterstalerstr. 103, 69118 Heidelberg eMail: n.groeben@uni-koeln.de

Haußecker, Nicole, geb. 1978. Wissenschaftliche Mitarbeiterin am Institut für Kommunikationswissenschaft der Friedrich-Schiller-Universität Jena.

Kontakt: Friedrich-Schiller-Universität Jena, Institut für Kommunikationswissenschaft Abteilung Kommunikationspsychologie, Ernst-Abbe-Platz 8, 07743 Jena eMail: nicole.haussecker@uni-jena.de

Hönig, Klaus, Dr. rer. nat., Dipl.-Psych., geb. 1970. Leitung der Konsil- und Liaisonpsychosomatik an der Universitätsklinik für Psychosomatische Medizin und Psychotherapie. Arbeitsschwerpunkte: Psychosomatik, Psychoonkologie, kognitive Neurowissenschaft, klinische Hypnose und Hypnotherapie, Leistungssport- und Business-Coaching.

Kontakt: Frauensteige 14a, 89075 Ulm eMail: Klaus.Hoenig@uniklinik-ulm.de

URL: http://www.uniklinik-ulm.de/index.php?id=10665

Jaeger, Susanne, Dr. rer. nat., geb. 1966. Wissenschaftliche Mitarbeiterin. Arbeitsschwerpunkt: Versorgungsforschung.

Kontakt: susanne.jaeger@zfp-zentrum.de

Jirschitzka, Jens, geb. 1979. Doktorand am Institut für Kommunikationswissenschaft der Friedrich-Schiller-Universität Jena und an der Doktorandenschule "Laboratorium Aufklärung".

Kontakt: Friedrich-Schiller-Universität Jena, Institut für Kommunikationswissenschaft Abteilung Kommunikationspsychologie, Ernst-Abbe-Platz 8, 07743 Jena eMail: jens.jirschitzka@uni-jena,

Keller, Ferdinand, Prof. Dr., geb. 1956. Leitender Psychologe in der Klinik für Kinder- und Jugendpsychiatrie am Universitätsklinikum Ulm. Arbeitsschwerpunkte: Statistische Methoden der Veränderungsmessung in Psychotherapie, Psychiatrie und Präventionsstudien, insbesondere hierarchische lineare Modelle und growth mixture models, Fragebogenentwicklung und psychometrische Analysen, insbesondere IRT-Verfahren (Rasch- bzw. Mixed-Rasch-Modelle, Latent-Class-Analyse), Depression, Behandlungszufriedenheit in der Kinder- und Jugendpsychiatrie, Evaluation von Therapieverfahren.

Kontakt: Klinik für Kinder- und Jugendpsychiatrie/Psychotherapie, Universitätsklinik Ulm Steinhövelstr. 5, 89075 Ulm

eMail: Ferdinand.Keller@uniklinik-ulm.de

Homepage Klinik: http://www.uniklinik-ulm.de/index.php?id=1420

Kiefer, Markus, PD, Dr. phil. habil., Dipl.-Psych., geb. 1966. Leiter der Sektion für Kognitive Elektrophysiologie der Klinik für Psychiatrie und Psychotherapie III der Universität Ulm, Privatdozent für Psychologie am Institut für Psychologie und Pädagogik der Universität Ulm. Arbeitsschwerpunkte: Kognitionspsychologie und ko-

gnitive Neurowissenschaft von Gedächtnis, Sprache Aufmerksamkeit und Bewusstsein.

Kontakt: Universität Ulm, Klinik für Psychiatrie und Psychotherapie III Sektion für Kognitive Elektrophysiologie, Leimgrubenweg 12, 89075 Ulm eMail: Markus.Kiefer@uni-ulm.de URL: http://www.uni-ulm.de/~mkiefer/

Markard, Morus, apl. Prof. Dr., geb. 1948. Derzeitige Tätigkeit: Evaluation & Forschungsmethoden und Praxisintegration in Pädagogischer Psychologie im Diplomstudiengang Psychologie an der FU Berlin. Arbeitsschwerpunkte: Theoretische und methodische Probleme einer subjektwissenschaftlichen Psychologie.

Kontakt: Fachbereich Erziehungswissenschaft und Psychologie der FU Berlin Studiengang Psychologie, Habelschwerdter Allee 45, 14195 Berlin eMail: mmarkard@zedat.fu-berlin.de

Messmer, Heinz, Dr. rer. soc. habil., geb. 1955. Professor an der Hochschule Soziale Arbeit, Fachhochschule Nordwestschweiz. Arbeitsschwerpunkte: Kerngebiete der Kinder- und Jugendhilfe, institutionelle Reaktionen auf abweichendes Verhalten Jugendlicher, konstruktivistische Konflikttheorie, ethnomethodologische Konversationsanalyse.

Kontakt: Fachhochschule Nordwestschweiz, Hochschule für Soziale Arbeit Institut Kinder- und Jugendhilfe, Thiersteinerallee 57, CH-4053 Basel eMail: heinz.messmer@fhnw.ch URL: http://www.fhnw.ch/sozialearbeit/ikj

Nohrstedt, Stig Arne, geb. 1946. Professor für Medien- und Kommunikationswissenschaften an der Universität Örebro, Schweden. Gastprofessor an der Linneaus Universität, Schweden. Arbeitsschwerpunkte: Kriegs- und Konfliktjournalismus, journalistische Ethik und Rollen, Risiko- und Krisenkommunikation, Medien und strukturelle Diskriminierung.

Kontakt: stig\_arne.nohrstedt@oru.se

Ottosen, Rune, geb. 1950. Professor für Journalistik am Oslo and Akershus University College for Applied Sciences. Arbeitsschwerpunkte: Geschichte des Journalismus, Kriegs- und Konfliktjournalismus.

Kontakt: Institute for journalism and media studies Oslo and Akershus University College of Applied Sciences, postboks 4 St. Olavspl., N-0130 Oslo eMail: Rune.Ottosen@hioa.no

Reetz, Klaus-Dieter, Dr. päd., geb. 1948. Betriebswirtschaftler und Diplom-Psychologe. Wissenschaftlicher Mitarbeiter des International Board of Economics and Business Knowledge e. V. (IBBK e. V.). Arbeitsschwerpunkte: Forschungsmethoden der Psychologie, ökonomische Bildung und Wissenschaftstheorien.

Kontakt: Eifelstrasse 12, 50677 Köln

eMail: mail@ibebk.org

Reimann, Michael, geb. 1966. Mediendidaktiker und Drehbuchautor für interaktive Lernmedien.

Kontakt: mr@michael-reimann.ch

Rost, Jürgen, Prof. Dr., geb. 1952. Seit 2005 i. R. Bis dahin wissenschaftlicher Direktor am Leibniz-Institut für die Pädagogik der Naturwissenschaften (IPN) an der Universität Kiel. Mitglied des PISA-Konsortiums 2003 und 2006. Arbeitschwerpunkte: Item-Response-Theorie, Latent-Class-Analyse, Mixed-Rasch-Modell, Umwelterziehung und nachhaltige Entwicklung.

Kontakt: an@j-rost.de URL: http://www.j-rost.de, http://www.zpid.de

Scheiblechner, Hartmann, Prof. Dr., geb. 1939, gest. 2010. 1972 bis 2004 Professor für Psychologie an der Philipps-Universität Marburg.

Hartmann Scheiblechner gehörte ebenso wie Wilhelm Kempf zur Wiener Schule der Messtheorie von Georg Rasch. Er distanzierte sich – ebenso wie Wilhelm Kempf – schon früh von der 'reinen Lehre' des Rasch-Modells, um neue psychometrische Wege zu gehen; wozu sich beide der konzeptionellen Grundlagen des Rasch-Modells bedienten. So untersuchte der eine die Bedingungen, unter denen Summenscores das Merkmal der spezifischen Objektivität aufweisen; der andere, inwieweit die Forderung nach stochastischer Unabhängigkeit durch eine bestimmte Art der stochastischen Abhängigkeit ersetzt werden kann.

Der hier abgedruckte Beitrag ist vermutlich die letzte Veröffentlichung von Hartmann Scheiblechner. Er starb am 24. Dezember 2010.

Spiess, Martin, Prof. Dr., geb. 1960. Professor für Psychologische Methoden an der Universität Hamburg. Arbeitsschwerpunkte: Schätzung nicht-linearer Modelle, Panelmodelle, Missing-Data-Techniken, Item-Response-Theorie, Survey-Statistik.

Kontakt: Psychologische Methoden, FB Psychologie, Universität Hamburg Von-Melle-Park 5, 20146 Hamburg URL: http://www.epb.uni-hamburg.de/de/node/817

Vollhardt, Johanna Ray, Ph.D., Dipl.-Psych., geb. 1979. Assistant Professor für Sozialpsychologie, Clark University. Arbeitsschwerpunkte: Friedenspsychologie, Genozid-Entstehung und -überwindung, Opferbewusstsein und Toleranz, Intergruppenbeziehungen zwischen Minderheitsgruppen.

Kontakt: Clark University, Psychology Department 950 Main Street, Worcester, MA 01610, USA eMail: JVollhardt@clarku.edu URL: http://vollhardt.socialpsychology.org/

Werbik, Hans, Prof. Dr., geb. 1941. Professor emeritus. Arbeitsschwerpunkte: Handlungspsychologie, Handlungstheorie, Wissenschaftstheorie.

Kontakt: 90562 Heroldsberg, Grossgeschaidt 306

# **Sachregister**

Auswertungsmodelle, statistische 11, 24ff.

BDI-II 120-126, 131f.
Bedingtheitsdiskurs s. Begründungsdiskurs
Begründungsdiskurs 14, 169ff.
Bedrohungserleben 43f., 53ff., 57, 59
Berichterstattung 11f., 40ff., 46, 48, 50, 52, 59-62, 68ff., 81, 83-91, 159
Berichterstattungsmuster (-stile) 12, 87ff., 91, 159
Bildgebung, funktionelle 190, 194f.

Depression 13, 120, 122, 127 Doppelbindung 61, 82 Dramatisierung 11, 42f., 48-52, 59

Experiment 11, 19, 23ff., 55, 95ff., 140-143, 150, 156, 165-168, 172, 186, 198

Fernsehnachrichten s. Nachrichten Frame 41ff., 50 Friedenskitsch 11f., 61f., 69-74, 80ff. Friedensjournalismus 9, 83f., 88, 90 Friedenspsychologie 11, 17, 23-26

Gedächtnis, semantisches 191, 193
Gegenstands-Methodik-Interaktion 11,
18, 23
generalized estimating equations (GEE)
12, 110-119
Gewaltspirale 26
Globalisierung 83f., 86, 185f.
Grounded Theory 71, 158
Gruppendiskussion 24, 160f., 179, 183186

Handlungsanalyse 179f., 184, 186 Handlungsfähigkeit 168 Handlungsforschung 143, 168, 179 / Entwicklungsfigur 173 Hirnforschung 197 Hypothese 13, 19-22, 26, 134-142, 149f., 158f., 164, 170f., 178, 181

Inhaltsanalyse 10, 68, 86, 91, 133, 139, 159, 184f.
Interaktionsanalyse 179, 184ff.
Isotonie 97
Itemparameter 96, 109, 111f., 114f., 123f., 126, 129-132
Item-Response-Modell 108f., 114

Kodiersystem 140
Konflikt als Prozess 11, 29f., 38
Konflikt als System 11, 29, 36
Konfliktemergenz 35
Konflikttheorie, konstruktivistische 11, 28, 30, 34
Konstruktivismus, methodologischer 10, 177f.

Konfliktberichterstattung s. Berichterstattung Kriegsberichterstattung s. Berichterstattung Kritische Psychologie 165f., 168

Latent-Class-Analyse 10, 12, 20, 25, 83, 86ff., 90f., 133, 141f., 149, 159, 184ff.

227 Sachregister

Maximum-Likelihood-Schätzer, marginaler (MML) 12f., 109, 114-119 Medienberichterstattung s. Berichterstattung

Medienframe s. Frame Mixed Methods 13, 154, 159f., 164 Mixed-Rasch-Modell 123f., 128, 131

Medien 11f., 41-44, 46, 48, 59-62., 82-90 Methoden (Forschung Psychologie)

qualitative

- quantitative

Modelle, mehrdimensionale 100-104 Moderatorvariable 145, 149 Muslime 12, 43, 56-59, 85

Nachrichten 11, 41, 44, 47-52, 58f., 61f., 86f., 91

Objektivität, spezifische 12, 95, 123, 132 Opferbewusstsein 22f. Ordinalskala 96f.

Partial-Credit-Modell 124, 127ff., 131
Peersozialisation 164
Plastizität, erfahrungsabhängige 192,
196
Populärkultur 61f.
Positivismus 136
Positivismusstreit 136, 155
Prämissen-Gründe-Zusammenhang 169174

Propaganda 12, 72, 85, 87, 89ff.

Pseudoempirie 171

Psychologie

- integrative 9, 14, 133, 150, 187f., 196-199

- subjektwissenschaftlich-interpretative 177ff.

Rasch-Modell 95, 98, 109, 114, 121-131, 142

Ratingskalen-Modell 124, 127ff. Reduktionismus, radikaler 197

Rekonstruktion, sozialpsychologische 150, 179f., 184

Rhetorik der Neuen Weltordnung 87

Rückkoppelungsgespräch, argumentatives 14, 176f., 180-185

Schätzung, nicht-iterative 113
Selbstkonzeptforschung 13, 160, 164
Simulationsexperiment 12, 110, 115-119
Stichprobenunterschiede 13, 123-132
Studentenbewegung 155
Subjektwissenschaft 165
Subskalenanalyse 121

Testmodell, probabilistisches 108, 120ff., 132, 142 Theorie-Praxis-Integration 11, 20

Umfrageforschung 166ff.

Validierung, kommunikative 14, 176, 182 Validität -externe 11, 18, 20, 23, 25 -interne 18-21, 29 Verallgemeinerung 171f. Verständnisbildung - intersubjektive 182, 186 - transsubjektive 183f., 186

Werturteilsfreiheit 17f., 25 WINMIRA 124

Ziel-Mittel-Argumentation 26



verlag irena regener berlin

Wir setzen Qualitätsstandards jenseits des Mainstreams

## Wilhelm Kempf (ed.)

fax: #49.30.42 85 67 62

eMail: verlag@regener-online.de

www.regener-online.de

# Readings in Peace Journalism Foundations – Studies – Perspectives

(Friedens- und Demokratiepsychologie, Bd. 9). 2010. 204 S., brosch., 41 Abb. & 22 Tab., € 29,90. ISBN 978-3-936014-22-8.

Since Galtung (1998) and Kempf (1996) outlined their first ideas of an alternative to conventional war reporting, their model(s) of "Peace Journalism" stimulated a broad debate among peace researchers and journalists, practical thought about how to achieve this type of journalism, and a large body of basic theoretical and empirical research. How the concept of peace journalism developed in the course of these studies is documented in the present book which contains a collection of papers from the years 1997–2009 that previously were only available in German, in hard to find sources and/or in electronic form.

#### Contents:

Part I, Foundations: War propaganda versus peace journalism (Wilhelm Kempf) / Conflict prevention and the media (Wilhelm Kempf) / Information structures as barriers to civilian conflict management (Wilhelm Kempf)

Part II, Studies: Media representations of conflicts as catalysts of peace processes (Susanne Jaeger & Wilhelm Kempf) / The German press coverage on France after World War II (Susanne Jaeger) / Acceptance and impact of de-escalation-oriented conflict coverage (Wilhelm Kempf) / Peace journalism and the news production process (Burkhard Bläsi)

Part III, Perspectives: Compatibility of peace and news media (Susanne Jaeger) / Media effects during violent conflict (Vladimir Bratić) / Implementing peace journalism: The role of conflict stages (Burkhard Bläsi) / The media and the implementation of a European peace policy (Wilhelm Kempf)

verlag irena regener berlin ostseestr. 109 10409 berlin



verlag irena regener berlin

Wir setzen Qualitätsstandards jenseits des Mainstreams

# Forschungsmethoden der Psychologie

Zwischen naturwissenschaftlichem Experiment und sozialwissenschaftlicher Hermeneutik

(Hochschullehrbücher, Bd. 2-4)

In konsequenter Auffassung der Psychologie als Wissenschaft von der subjektiven Welt des Menschen, deren Gegenstandsbereich von den physiologischen Ursachen bis hin zu den Prozessen der sozialen Konstruktion von Wirklichkeit reicht, verfolgen die Autoren eine wissenschaftstheoretische und didaktische Konzeption, welche die Methoden der Psychologie im Spannungsfeld zwischen naturwissenschaftlichem Experiment und sozialwissenschaftlicher Hermeneutik verortet, eine gegenstandsangemessene Umsetzung des galileischen Denkens in der Psychologie zum Ziel hat und die Kluft zwischen natur- und kulturwissenschaftlicher Herangehensweise an den Forschungsgegenstand der Psychologie ebenso zu überwinden trachtet wie jene zwischen quantitativen und qualitativen Forschungsmethoden.

#### Band I:

Wilhelm Kempf, Theorie und Empirie

3. Auflage. 2009. 377 S., 377 S., brosch., mit 71 Abb. und 34 Tab., 42,90 €. ISBN 978-3-936014-21-1.

#### Band II:

Wilhelm Kempf, Quantität und Qualität

1. Auflage. 2008. 396 S., 396 S., brosch., mit 47 Abb. und 59 Tab., 44,90 €. ISBN 978-3-936014-15-0.

#### Band III:

Wilhelm Kempf & Markus Kiefer (Hrsg.), Natur und Kultur

1. Auflage. 2009. 366 S., 366 S., brosch., mit 63 Abb. und 8 Tab., 44,50 €. ISBN 978-3-936014-17-4.

Alle drei Bände im Paket zum Sonderpreis: Statt 132,30 € nur 120,- €! ISBN 978-3-936014-20-4.

verlag irena regener berlin ostseestr. 109

10409 berlin

fax: #49.30.42 85 67 62 www.regener-online.de eMail: verlag@regener-online.de