

Ludwig, Peter H.

Bewertungskriterien für hypothesenprüfende qualitative Bildungsforschung: keine, alternative oder klassische Qualitätsstandards?

2010, 30 S.

urn:nbn:de:0111-opus-30696



Nutzungsbedingungen / conditions of use

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

We grant a non-exclusive, non-transferable, individual and limited right to using this document.
This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.
By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
Deutsches Institut für Internationale Pädagogische Forschung (DIPF)
Mitglied der Leibniz-Gemeinschaft
Informationszentrum (IZ) Bildung
Schloßstr. 29, D-60486 Frankfurt am Main
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Peter H. Ludwig

Universität Koblenz-Landau: Institut für Bildung im Kindes- und Jugendalter

Bewertungskriterien für hypothesenprüfende qualitative Bildungsforschung: keine, alternative oder klassische Qualitätsstandards?

Überarbeitete Fassung eines Vortrags und präsentierten Papers auf dem Kongress „Contested Qualities of Educational Research“ der „European Educational Research Association“ (EERA) an der Universität Gent, Belgien, am 21. Sept. 2007 [‘Anything goes – quality stays’. Standards of quality for qualitative empirical research in education]

Zusammenfassung: Auch qualitative Untersuchungen zu Lern- und Bildungsprozessen weisen qualitative Unterschiede auf. Um die unterschiedliche Aussagekraft von qualitativen Studien im Begründungszusammenhang bestimmen zu können, ist es erforderlich, sich – ähnlich wie bei quantitativer Forschung – auf Standards zu einigen. Es wird die Auffassung begründet, dass eine Laissez-Faire-Haltung oder das Aufstellen ersetzender „alternativer“ Wissenschaftlichkeitskriterien wenig zum Ansehen qualitativer Ansätze beizutragen vermag. Die umfassende Etablierung qualitativer Verfahren zur Gewinnung belastbarer Erkenntnisse hängt entscheidend von der Anerkennung und Berücksichtigung konsensfähiger allgemeiner Qualitätskriterien für empirisches Vorgehen ab. Es wird dafür plädiert, dass qualitative Ansätze als empirische Verfahren prinzipiell denselben wissenschaftlichen Regeln unterliegen sollten wie quantitative.

Abstract: Quantitative as well as qualitative research studies on topics of education within the context of justification vary in their level of quality. However, unlike researchers working with quantitative empirical methods, not all of their qualitatively oriented colleagues do recognize a mutual set of quality standards for their studies. In order to determine the degree to which we can reliably learn from a qualitative investigation, it is necessary to reach agreement on standards of quality within the qualitative paradigm. Neither a laissez-faire attitude of „anything goes“ nor the attempt to create „alternative“ standards to the „traditional“ ones do contribute to the reputation of the qualitative approach. Its extensive establishment as indispensable research methods of obtaining sound information depends on the acceptance and use of agreed-upon general criteria for determining the quality for methodological procedures.

Bereits etwa zehn Jahre nach dem Ausruf der „realistischen Wende“ 1962 durch Heinrich Roths vielbeachtete Göttinger Antrittsvorlesung setzte in der erziehungswissenschaftlichen Forschung parallel zum Einsatz quantitativer

Verfahren ein zunehmend starker Trend zu qualitativen Methoden ein. Der („neue“) qualitativ-interpretative und der („traditionelle“) quantitative Methodenbereich galten lange Zeit als weitgehend unvereinbar. Ihre jeweiligen Vorzüge und Schwächen wurden von ihren Protagonisten und Antagonisten gern überspitzt diametral gegenübergestellt. Dieser Disputationsstil in Form eines „methodologischen Grabenkrieges“ (Frey 1983, 279; Flick 2002, 381; 2004, 67) wurde von den qualitativen Methodologen insofern mit angefeuert, als sie ihre Ansätze oftmals in schroffem Kontrast zur quantitativen („konventionell-positivistischen“) Methodik formulierten und diese zu ihrem „Lieblingsfeind ... erklärten“ (Kalthoff u.a. 2008, 87; vgl. Baumann 1998, 140; Banister et al. 1994, iii, 1) und dabei teilweise in Kauf nahmen, sich von einem Zerrbild als von der Realität quantitativer Forschung abzugrenzen. Einige quantitativ orientierte Forscher hingegen weigerten sich, qualitative Methoden überhaupt zu den *empirischen* Verfahren zu zählen. Die ausgetauschten Argumente beider Seiten offenbarten – neben berechtigten Einwänden – jedenfalls nicht selten, dass die Kontrahenten von der attackierten Richtung eine eher begrenzte Vorstellung zu besitzen schienen (Saldern 1995, 340). Gängigen Argumentationsstereotypen zufolge würden quantitative Ansätze der Illusion gesicherter Erträge unterliegen, während qualitative näher am Menschen, am Alltag und der pädagogischen Praxis seien (Ludwig 2003).

Die ursprüngliche Konfrontation zwischen beiden Forschungsmethodenbereichen ist inzwischen einer beidseitig behutsamen Annäherung gewichen. Folgende Indikatoren machen diese Tendenz sichtbar:

(1) Ideologische Schranken fallen: Gemeinsamkeiten werden stärker herausgestellt als prinzipielle Unterschiede. Zunehmend verbreitet sich die Einsicht, dass mit einer scharfen Distinktion nur eine Scheinkontroverse geführt wird (Früh 1991, 230ff; Mayring 1999, 292; Helmke/Schrader 2006, 92; vgl. Ludwig 2004; Menck 2005, 10). Da es sich nicht um „Glaubensrichtungen“ handle, bestehe keine Notwendigkeit, sich alternativ und grundsätzlich einer der beiden Methodenrichtungen entweder anzuschließen oder zu verweigern (Niegemann 2006, 390f). Vielmehr bestimmen der jeweilige Forschungsgegenstand und der bislang erreichte Erkenntnisstand die Eignung einer Forschungsmethode. Grundsätzliche Methodenvorlieben würden das Sichtfeld auf bestimmte Fragestellungen unnötig und kontraproduktiv eingrenzen (Danner 2006, 16, 19; Niegemann 2006, 390; Roeder 2006, 641).

(2) Auch quantitativ orientierte Fachvertreter plädieren inzwischen für den Einsatz qualitativer Methoden (z.B. Klieme/Bos 2000; Renkl 1999; Roeder 2006) und umgekehrt (z.B. Flick 2004, 67ff). In allgemein einführenden Lehr-

büchern zur (quantitativen) Forschungsmethodik werden inzwischen auch qualitative Verfahren erörtert (z.B. Bortz/Döring 2006; Diekmann 2001, 443ff, 510ff; Klauer 2001, 83ff; Cohen/Manion/Morrison 2007).

(3) Die Vereinbarkeit und Komplementarität beider Methodenbereiche werden hervorgehoben (z.B. Hammersley 1996, 167f; Abel/Möller/Treumann 1998, 154ff; Creswell 2003; Eye 1994, 36f; Ludwig 1999, 229f; Schnaitmann 2004; Kelle 2007; Finkbeiner u.a. 2006; Flick 2002, 380ff; Wellenreuther 2000, 13; Groeben 2006, 9). Mit Denzin (1970; vgl. Schwartz-Shea 2009, 26) fällt die pluralistische Integration dieser *zwei* Methodenbereiche unter den aus der Feldmessung entlehnten (mehrdeutigen und „numerisch“ gewöhnungsbedürftigen) Terminus „methodische *Triangulation*“.

Der gegenwärtige methodologische „Zeitgeist“ kommt also der qualitativen Methodik entgegen. Eine gewisse Etablierung der qualitativen Forschung ist in der Erziehungswissenschaft (weniger ausgeprägt in der Pädagogischen Psychologie) bereits erreicht. Wernet sieht darin gar eine „zweite realistische Wende“ der Erziehungswissenschaft (2006). Trotz ihrer nicht unbedeutlichen Verbreitung kommt der qualitativen Richtung - auch nach Ansicht qualitativer Methodologen - in vielen pädagogischen Forschungsfeldern dennoch eine bislang nur randständige Bedeutung zu (z.B. Blömeke 2004, 63; Leutner 2000, 32, 35; Reichertz 2000; Wellenreuther 2005, 53). Groeben bezeichnet die quantitative Methodik für die Psychologie als den *Mainstream* und die qualitative als den *Offstream*. Er spricht sogar provokativ von der „selbstverschuldeten Irrelevanz des qualitativen Forschungsansatzes“ (2006). Zudem kann von einer uneingeschränkten Akzeptanz jeder Form qualitativer Methodik derzeit noch nicht die Rede sein (Groeben 2006, 4; Lüders 2000, 635; Reicher 2000, 2; Flick/Kardorff/Steinke 2000, 13). Oft wird eine vermeintlich breite Anerkennung qualitativer Forschung mit der Existenz eigener Publikationsorgane, regelmäßiger Fachtagungen und der personalen Vertretung auf Hochschullehrer-Ebene belegt (z.B. Knoblauch 2008, 210). Solche Strukturmerkmale sind in der Tat Indikatoren für eine gewisse Größe des Zugehörigen-Kreises, jedoch nicht zwangsläufig für eine vorbehaltlose Anerkennung *außerhalb* dieses Kreises (vgl. Helsper/Herwartz-Emden/Terhart 2001, 252).

Zur Erlangung der Wertschätzung der gesamten Wissenschaftsgemeinde gegenüber den qualitativen Verfahren ist es erforderlich, noch Verständigung darüber herbeizuführen,

(1) was genau unter qualitativer bzw. quantitativer Forschung zu verstehen ist (vgl. Knoblauch 2008, 211ff; Ludwig 2004; Reichertz 2007),

- (2) wie eine „gemeinsame metatheoretische Basis“ (Klieme/Bos 2000, 361) aussehen könnte und vor allem
- (3) welche gemeinsam getragenen Qualitätsstandards oder Gütemaßstäbe gelten sollen.

Mit den folgenden Überlegungen zum dritten Punkt soll die Notwendigkeit der Einigung auf gemeinsame Maßstäbe unterstrichen werden. Dabei wird pauschalisierend von *der* qualitativen Forschung gesprochen. Differenzen zwischen den einzelnen qualitativen Ansätzen bleiben aus thematischen und räumlichen Gründen also weitgehend ausgeblendet, wohl wissend, dass dies der Mannigfaltigkeit der Ansätze nicht gerecht wird (Hirschauer 2008, 179). Diese einzelmethoden-übergreifende Sicht rechtfertigt sich erstens durch eine entsprechende verbreitete Denkgewohnheit, die pragmatischen Zwecken der Verständigung durchaus entgegen kommt (z.B. auch Helfferich 2004), zweitens durch die Existenz des Dachbegriffs „qualitativ“ selbst, der sinnleer wäre, wenn nicht von einem Minimum gemeinsamer konstituierender Merkmale qualitativer Ansätze ausgegangen werden dürfte, und drittens durch die Aussichtslosigkeit, die große Fülle an inzwischen vorgelegten unterscheidbaren, aber auch wenig unterscheidbaren Verfahren systematisch zu überblicken (Hirschauer 2008, 179).

1. Standpunkte zum Umgang mit Standards

Im Rahmen der erfahrungswissenschaftlichen Forschungs- und Methodologietradition wurde eine Reihe von Wissenschaftlichkeitskriterien erarbeitet. Dazu gehört, dass wissenschaftliches Denken und Argumentieren systematisch methodisch vorgeht (etwa auch Alternativ-Thesen berücksichtigend), nachvollziehbar, d.h. kommunizierbar, prüfbar, kontrollierbar ist, begründet wird, erkennbare Trennungen zwischen Ergebnissen und Interpretationen vornimmt und als solche bezeichnet, Fehlerquellen und Defizite offenlegt und sich an den Idealen von Objektivität und Verallgemeinerbarkeit orientiert (Klieme/Bos 2000, 376; Roeder 2006; Breuer/Reichertz 2001, 4; Terhart 1995, 375f; Kerlinger 2000; Rost 2009, 76). Derartige Kriterien geben in forschungsmethodischer Hinsicht Qualitätsstandards und damit einen Maßstab zur Beurteilung der Güte empirischer Studien vor, die unter quantitativ Forschenden als Leitideen weitgehend unumstritten sind, in der Forschungspraxis berücksichtigt werden und Gegenstand zumindest punktueller Überprüfungen sind.

Unter den Forschenden, welche sich selbst dem qualitativen Paradigma zurechnen, ist hingegen bis heute noch keine umfassende Einigung auf diese

oder andere Standards abzusehen (Flick 2004, 100; Lüders 2003, 84; Combe u.a. 1999; Brühl/Buch 2006, 7). Qualitative Forschung, soweit sie als Gegenentwurf zur quantitativen konzeptualisiert wird und damit dieser „ihre Identität verdankt“ (Kalthoff u.a. 2008, 87), tritt bisweilen mit dem Anspruch an, einengend empfundene Kriterien traditionell-konservativer Empirie aufzubrechen und progressive Wege zu erproben. Einem solchen emanzipatorischen Geist mag es naturgemäß zunächst wenig entsprechen, wiederum (neue) Reglementierungsdiskussionen zu führen (vgl. Helsper/Herwartz-Emden/Terhart 2001, 252). Möglicherweise liegt darin der Grund, weshalb die Thematisierung von Qualitätsstandards qualitativer Forschung, die über den Bezugsrahmen einzelner Methoden und Verfechter hinaus für die gesamte qualitative Richtung gelten sollen, innerhalb der eigenen Reihen teilweise noch tabuisiert wird (Steinke 1999, 10-12; vgl. Mayring 1999, 293). Dennoch ist die Verständigung auf Qualitätsstandards für Wissenschaft essentiell. Zur Ankurbelung dieser Debatte widmete die Zeitschrift für Erziehungswissenschaft im Jahr 2000 ihr Heft 3 dem Themenschwerpunkt „Standards qualitativer Forschung“, wobei die meisten Beiträge auf Qualitätsstandards jedoch explizit nur beiläufig eingingen. Auch im Online-Journal „Forum Qualitative Sozialforschung“ ist seit 2000 eine heftübergreifende Beitragsreihe zu dieser Debatte angelegt (Breuer u.a. o.J.).

In der qualitativen Forschung werden drei Positionen zu Qualitätsstandards vertreten (Flick 2002, 319; Saldern 1995, 354; Seale 1999; Steinke 1999, 50f; 2000, 319-321; Brühl/Buch 2006, 8; Smith/Hodkinson 2005; Schwartz-Shea 2009):

- die Zurückweisung jeglicher Standards, oft postmodern-konstruktivistisch untermauert
- die Aufstellung eigener, „alternativer“ Standards unter Ablehnung konventioneller Standards
- die Anerkennung derselben („traditionellen“) Standards, wie sie der quantitativen Forschung zugrunde liegen und dort allgemein akzeptiert werden

Die Stichhaltigkeit dieser drei Standpunkte soll im Folgenden geprüft werden. Dabei wird eine normative Position, d.h. eine bewertende Stellungnahme zu diesen drei Standpunkten abgegeben, begründet und zur Diskussion gestellt. Es geht also nicht um eine Fortsetzung des Diskurses um qualitative Methoden, sondern um die meta-methodologische Ebene der ihnen zugrundeliegenden regulativen Entwürfe.

Im Rahmen der Diskussion um Forschungsstandards ist die erkenntnistheoretische Einteilung von Forschungsphasen nach Reichenbach (1938, 7, 382; 1953, 260) in den „Entdeckungszusammenhang“ (context of discovery) und

den „Begründungs“- oder „Rechtfertigungszusammenhang“ (context of justification) wesentlich. Übertragen auf die empirische Forschung umfasst die erste Phase explorative Erkundungsuntersuchungen mit dem Ziel, Hypothesen- und Fragestellungen zu generieren bzw. Theorieentwicklung und –differenzierung oder Modellbildung zu betreiben. Im Begründungskontext bewegen sich hingegen hypothesenprüfende explanative Studien bzw. Untersuchungen, welche belastbare Antworten auf Forschungsfragen anstreben.

Die folgenden Überlegungen zu Qualitätsstandards beziehen sich in erster Linie auf die Forschungsphase des Begründungszusammenhangs. Der Entdeckungskontext als eine im Wesentlichen intuitive Domäne hingegen unterliegt weit gemäßigteren Anforderungen. Diese erste Forschungsphase muss keinem strengen Reglement genügen. Hypothesen können auch in einem kreativen Prozess gewonnen werden, „für den es keine objektiven Regeln gibt“ (Beutelspacher 2002, 65).

Qualitativen Verfahren wird häufig die Hypothesengenerierung zugesprochen und quantitativen die Hypothesentestung (z.B. Roeder 2006, 640f; Flick 2002, 382; Klauer 2001; Bohnsack 2008, 13ff; Brühl/Buch 2006, 1). Eine der Stärken der qualitativen Forschung mag zwar in der Tat in der Beisteuerung gründlicher Analysen des Felds zur Hypothesengewinnung liegen; dennoch entspricht diese Limitierung auf „Vorstudien“ nicht dem Selbstverständnis vieler qualitativer Forscher und ihrer Projekte (z. B. Hopf 1996; König/Zedler 1995). Dies ist begreiflich: Denn so bedeutungsvoll die *Entwicklung* von „guten“ Hypothesen bzw. „lohnenden“ Fragen sein mag; die gesellschaftliche Funktion der empirischen Erziehungswissenschaft erschöpft sich letztlich nicht in der Aufwerfung bedeutsamer Fragen oder in der Aufstellung plausibler Vermutungen – dies gelingt engagierten Praktikern oder theoretisch Forschenden bisweilen auch zufriedenstellend -, sondern zielt primär auf deren zutreffende verteidigungsfähige Beantwortung, weil im Begründungszusammenhang erfahrungswissenschaftliches Know-How kaum durch bloße Kreativität oder Rationalität – wie im Entdeckungszusammenhang - ersetzbar ist (vgl. Lamnek 1995, 223ff; Lodge 1978, 45). Dieser Erwartung durch eine „evidence-based reform in education“ stärker als bislang zu entsprechen, fordert etwa Slavin (2002; 2007; 2008; Bridges 2008; Keiner 2008; Jornitz 2008). Für qualitative Untersuchungen, die sich jedoch im Begründungszusammenhang bewegen und damit den Anspruch erheben, Theorien oder Hypothesen zu prüfen, treffen die folgenden Überlegungen zu den Kriterien wissenschaftlicher Qualität zu (Brühl/Buch 2006, 1).

2. Begründung des Erfordernisses von Qualitätsstandards

Die völlige Ablehnung von Standards für Studien im Begründungskontext ist näher betrachtet eine kaum haltbare Position. Freilich ist Wissenschaft dem modernen Verständnis nach grundsätzlich ein offenes Unternehmen (Popper 1966, 223ff; 1992). Nicht nur der Pool wissenschaftlich produzierter Erkenntnis verändert sich durch Ausbau oder Korrektur. Auch das Wissenschafts(selbst)verständnis ist nicht „in Stein gehauen“, sondern grundsätzlich diskutierbar und unterliegt tatsächlich einem bedächtigen historischen Wandel. Gleichwohl können sich Wissenschaft im Ganzen, aber auch Einzelwissenschaften innerhalb eines epochalen Abschnitts kaum anders als über einen Satz anerkannter „Spielregeln“ konstituieren, um nicht zu einem beliebig austauschbaren und damit zweckfreien „Glasperlenspiel“ sensu Hesse zu werden (Prein/Erzberger 2000, 346; Hesse 1946; Kelle/Kluge 2006).

Kritiker mögen formulierte Regeln wissenschaftlichen Arbeitens als künstliches Korsett empfinden (vgl. Hirschauer 2008, 184). Für eine solche „Methodentoleranz“ stehen etwa der methodologische Postmodernismus und die anarchistische Wissenschaftsphilosophie Paul Feyerabends sowie in der Pädagogik der Ansatz Heinrich Rombachs (1965). Feyerabend wollte mit seinem libertär-methodologischen Schlachtruf „anything goes“, der einem gleichnamigen Musical von Cole Porter (1934) entliehen ist, Wissenschaft vom Joch einer vermeintlich kleinlichen Auswahl zulässiger Methoden befreien (1979; 1983; Feyerabend/Thomas 1986; vgl. Bamme 1986). (Er dachte dabei zwar eher an Natur- und Formalwissenschaften; dennoch scheinen sich Vertreter von geistes-, kultur- und sozialwissenschaftlichen Disziplinen von solchen Maximen stärker angezogen zu fühlen.) Methodische Freiheit ist jedoch nur innerhalb gewisser Grenzen erkenntnisförderlich. Denn von überzogener Offenheit ist es nur ein Schritt zur inhaltlichen Unverbindlichkeit und damit zur Beliebigkeit, Willkürlichkeit und Belanglosigkeit; letztlich bis hin zur Bewahrheitung des scherzhaft gemeinten Aphorismus, der jeweilige Stand der Wissenschaft verkörpere lediglich den ‚augenblicklich geltenden Irrtum‘. Der Slogan *anything goes* wurde offenbar von Beipflichtenden ernster genommen als er vom Protagonisten gemeint war (Feyerabend 1995, 196, 221, 229).

Diejenigen, welche für eine solche totale Offenheit plädieren, manövrieren sich in die Schwierigkeit, die Identität wissenschaftlichen Denkens bestimmen und damit wissenschaftliches Erkenntnisstreben von außer- oder vorwissenschaftlichen Produktionen abheben zu können, wie sie etwa von pädagogischen Praktikern, Erziehungsratgebern, Feuilletonisten und Essayisten hervorgebracht werden. Wissenschaft verlöre damit ihre gesellschaftliche

Funktion, nämlich Hort des prinzipiell zustimmungsfähigen Wissensbestands zu sein. Dieser Erkenntnispool ist zwar vom Stand der Forschung abhängig und damit dynamisch, aber dennoch aktuell verbindlich, da er jeweils zeitbezogen die *bestmögliche* Basis für öffentliche Verständigungen und Entscheidungen darstellt.

Pädagogische Vermutungen, Spekulationen oder Hypothesen können auch von Laien formuliert werden. Der einzige, aber entscheidende Vorzug, den hingegen empirisch geprüfte erziehungswissenschaftliche Aussagen haben und der den unvergleichlich höheren Aufwand bei der Produktion dieses Wissens rechtfertigt, ist ihr höherer Grad an Gesicherheit. Wem es nicht auf möglichst belastbare Aussagen ankommt, der hat wenig Interesse an Wissenschaft. Daher ist diese Zielsetzung –unabhängig von der Forschungsmethodik – ein Gebot, das nicht in Frage gestellt werden kann, ohne die Wissenschaftlichkeit selbst zur Disposition zu stellen. Je stärker subjektiv-beliebige Einflüsse bei der Untersuchung einer Forschungsfrage hineinspielen, desto geringer wird die Belastbarkeit ihrer Befunde sein. Dieses Grunddilemma, das durch das disproportionale Verhältnis von Sicherheit und Subjektivität entsteht, dürfte kaum auflösbar sein.

Die Frage, was Wissenschaft ist oder sein soll, lässt sich nur aufgrund der Regeln wissenschaftlichen Denkens beantworten. Das Einvernehmen hinsichtlich eines gewissen Mindestbestands an solchen Regeln könnte als der „Reifegrad“ einer Disziplin bezeichnet werden. Standards geben einen Orientierungs- und Bewertungsmaßstab ab für die Legitimation, Durchführung und Einschätzung von Forschung (Steinke 1999, 11, 2000, 321). Grundsätzlich kann zwar frei entschieden werden, wie hart oder weich wissenschaftliche Standards gefasst werden sollen; bei den Konsequenzen dieser Entscheidung besteht jedoch kein Freiheitsgrad mehr: Je höher Standards gesetzt werden, desto verlässlicher sind die Erträge der Forschung, welche diese Standards einhalten, und desto größer dürfte die gesellschaftliche Relevanz einer Disziplin und damit (berechtigterweise) ihre öffentliche Reputation sein. Je offener und weicher ihre angelegten Kriterien gehalten sind, desto beliebiger, austauschbarer und bedeutungsärmer werden ihre Befunde. Forschungszweige, die allzu leger mit Standards umgehen, partizipieren allenfalls (unverdient) mit am Kapital des öffentlichen Vertrauens auf wissenschaftlichen Erkenntnissen, das Untersuchungen erwirtschaftet haben, welche strengeren Standards genügen.

Ein allzu unverbindlicher Umgang mit Forschungsstandards wird von außen leicht als Selbstgefälligkeit empfunden. Nur mit einer gehörigen Portion Zynismus ließe sich so manche strittige methodologische Argumentation etwa

auf die medizinische Forschung übertragen. (Diese ebenfalls human- und handlungswissenschaftliche Disziplin ist soweit durchaus mit der Erziehungswissenschaft vergleichbar; siehe Reinmann 2006, 17.) Der Aufruf zur methodischen Unverbindlichkeit mag in sozial- und verhaltenswissenschaftlichen Kontexten nur deswegen leichter fallen, weil sich die praktischen Folgen von Fehlaussagen dort zumeist weniger prompt einstellen und damit im Einzelfall weniger eindeutig auf konkretes Handeln rückführbar sind. Dies entbindet die Forschung keineswegs von ihrer Verantwortung für ihre Erträge.

3. Einwendungen gegen „alternative“ Standards

Die Befürworter eigener „qualitativer“ Standards sehen in der Anerkennung konventioneller Regeln mitunter eine Unterwerfung unter eine vermeintliche Vorherrschaft quantitativer Methodik (z.B. Kardorff 1995, 4; vgl. Groeben 2006, 8). Es gibt jedoch keine „*quantitativen* Standards“, sondern nur solche, welche von der quantitativen Forschungsrichtung *anerkannt* werden. Lediglich die Verfahren zur Einhaltungsüberprüfung sind teilweise quantitativer Natur, etwa die Einschätzung der Interraterobjektivität mit Hilfe von Korrelationskoeffizienten.

Die üblichen Standards beruhen auf richtungsübergreifenden, rational gut begründeten methodologischen Argumenten. Sie fußen teils auf umfangreichen Erfahrungen mit (häufig in meta-experimentellen Untersuchungen) nachgewiesenen Fehlerquellen und Forschungsartefakten, mit denen in der qualitativen Forschung nicht minder zu rechnen ist; z.B. Vorurteils- und Stereotypeneinflüsse, Effekte sozialer Erwünschtheit, Interviewer-Einflüsse, Beurteilungsfehler, Einflüsse der Frageformulierung, Reaktivität von Erhebungsinstrumenten oder selektive Wahrnehmung. Die arglose Ignoranz gegenüber einem solchen Störpotential entspräche einer Kopf-in-den-Sand-Illusion, da derartige Fehlerquellen durch Nichtbeachtung nicht zu eliminieren sind.

Die Haltbarkeit des Rufs nach alternativen Standards soll im Folgenden exemplarisch anhand einer zentralen Ausdifferenzierung von Qualitätsstandards erörtert werden, nämlich den drei klassischen messtheoretischen Gütekriterien: Objektivität, Reliabilität und Validität (Lamnek 2005, 138-179; Moosburger/Kelava 2007). Steinke etwa lehnt für qualitative Forschung Gütekriterien partiell ab. Sie fordert die Entwicklung „spezifischer Bewertungskriterien“, die letztlich aber wiederum keine Allgemeingültigkeit beanspruchen, sondern projektspezifisch zur Anwendung kommen sollen (1999, 204f, 249, 251). Teils werden auch „methodenangemessene Gütekriterien“ als Ersatz der klas-

sischen konzipiert (Flick 2004, 100; Schwartz-Shea 2009, 33ff; Helsper/Herwartz-Emden/Terhart 2001, 258).

Dem ist entgegen zu halten: Essentielle, wissenschaftskonstitutive Kriterien dürften kaum „ersetzbar“ oder „methodenangemessen“ spezifizierbar sein. Standards als die „Verfassung“ von Forschung sind einzelnen Forschungsverfahren vor- und nicht nachgängig. Daher kann die Argumentation Knoblauchs (2000, 628), deswegen für die Entwicklung eigener Kriterien zu plädieren, weil mit dem Anlegen der herkömmlichen Kriterien der „Verlust des spezifischen Beitrags“ der qualitativen Forschung zu befürchten sei, kaum überzeugen. Denn sie stellt die Prioritätslogik von Forschungskriterien und Forschungsmethodik auf den Kopf.

Ein *Ersatz* käme in seiner Dramatik der Weigerung eines Pharma-Unternehmens gleich, Verbindliches zu den (Neben)Wirkungen seiner Produkte auszuweisen, „alternativ“ dafür jedoch zu versichern, dass die Produktinformation auf Recycling-Papier gedruckt wurde. *Methodenangemessene* Ausgestaltung würde analog etwa im Straßenverkehrsrecht bedeuten, den Alkohol-Grenzwert am Steuer nicht an der notwendigen Fahrtüchtigkeit und öffentlichen Sicherheit auszurichten, sondern variabel den individuellen Trinkgepflogenheiten deswegen anzupassen, weil es ansonsten einigen Verkehrsteilnehmern schwer fallen würde, sich innerhalb der Legalität zu bewegen.

Die gemeinsame Anerkennung eines gewissen Pools von Forschungsstandards ist nach Brühl & Buch (2006) die Voraussetzung für eine Vereinbarkeit von quantitativen und qualitativen Methoden auf wissenschaftstheoretischer Ebene. Kriterien zur Beurteilung von Forschungsqualität implizieren im Grunde per se ihre *pauschale* Anwendung, zumindest innerhalb einer Disziplin. Entgegen anderer Auffassungen (z. B. Hirschauer 2008, 178, 184) wird hier also die Position einer weitgehend einheitswissenschaftlichen Methodologie vertreten. Denn es gehört zum Wesen von Standards, mit einem gewissen Universalitätsanspruch ausgestattet zu sein. Im Gegensatz zu esoterischen (d.h. „nach innen gewandten“) Lehrgebäuden ist wissenschaftliche Erkenntnisproduktion darauf ausgerichtet, jenseits des Kreises der unmittelbaren Produzenten Anerkennung zu finden. Damit werden auch die anzulegenden Gütemaßstäbe der Forschung auf *allgemeine* Billigung ausgerichtet sein müssen. Dies schließt eine nur sich selbst genügende methodische Ausrichtung aus. Denn wie wäre zu rechtfertigen, dass unterschiedliche Auffassungen, etwa gegenüber der Forderung nach Intersubjektivität und Nachvollziehbarkeit von Forschungserträgen oder nach Offenlegung von Fehlerquellen, je nach methodologischer Orientierung der Forschenden in ein und

derselben wissenschaftlichen Disziplin koexistieren (sollen)? Wenn sich jeder methodische Ansatz durch sich selbst legitimieren könnte, wäre der Wissenschaftsbegriff ungehindert auf jede denkbare Form menschlicher Erkenntnis anwendbar (Brezinka 1989, 144, 146). Daher sind „Privatkriterien“, die nur für einzelne Forschergruppen oder Methoden gelten sollen, keine „Standards“, ähnlich wie Rechtsfestlegungen nicht ad personam zugeschnitten sein können. Dennoch mögen spezielle Übereinkünfte als Ausweis eines individuellen Qualitätsanspruchs im Sinne einer *zusätzlichen* Selbstverpflichtung nützlich sein. Die Verständigung auf einen Katalog von methodenübergreifenden Mindeststandards schließt also dessen Ergänzung durch methodenspezifische Maßstäbe nicht aus, durch die die besonderen Leistungen und Möglichkeiten einzelner „Vorgehensweisen sichtbar werden können“ (Kardorff 1995, 4). Metatheoretische Postulate, welche einen besonderen Schonraum oder Sonderrechte für qualitative Forschung reklamieren, indem sie sich auf alternative spezifische Bewertungskriterien berufen, präsentieren diese Methodik indirekt weit unter Wert. Kurioserweise werden als vermeintlich „alternative“ bzw. „methodenspezifische“ Standards auch solche vorgeschlagen, welche sich mit konventionellen Vorstellungen ohnehin decken, wie z.B. die Zuverlässigkeit, Nachvollziehbarkeit, Reproduzierbarkeit, die Dokumentation oder Regelgeleitetheit des methodischen Vorgehens (z.B. Flick 2002, 333f, 343; Lüders 2000, 633; Esser 1987, 88).

Objektivität, Reliabilität und Validität sind eine zurückhaltende (operationalisierungsfähige) Umschreibung für das „Wahrheitskriterium“ in Bezug auf Erhebungsinstrumente (vgl. Helfferich 2004). In den Realwissenschaften können sie deshalb als essentiell gelten: „Ob eine Aussage als wissenschaftlich akzeptiert wird, hängt davon ab, ob der Erkenntnisweg objektiv, zuverlässig und gültig ist“ (Saldern 1995, 353; vgl. Hopf/Müller 1995, 67; Mayring 1999, 297). Wer diese Gütekriterien grundsätzlich und radikal zurückweist, verabschiedet sich vom empirischen Denken. Denn dies hieße, es für belanglos zu halten, ob das, was erhoben werden soll, auch de facto erhoben wurde. Es hieße zu billigen, dass Merkmale beliebig ungenau erfasst werden können; und es mit einem Achselzucken zu kommentieren, wenn Sekundäranalysen oder Replikationsstudien völlig andere Ergebnisse zu Tage fördern würden. Solche Gütekriterien lassen sich nicht umfassend ablehnen, ohne gleichzeitig die prägnanten Konturen zwischen Wissenschaft und nicht-wissenschaftlichen Aussagesystemen und Ideologien zu verwischen.

Der grundsätzlichen Wertschätzung der Gütekriterien widerspricht es jedoch keineswegs,

- dass für neuartige Erhebungsmethoden auch *neue Prüftechniken* - also neue Operationalisierungen - der klassischen Gütekriterien entwickelt

werden (Diekmann 2001, 452; Flick 2002, 319ff; Bortz/Döring 2006, 326ff), etwa die „kommunikative“ bzw. „dialog-konsenstheoretische Validierung“ (Sogar einige der gebräuchlichen Operationalisierungen der Gütekriterien, z.B. die Splithalf- und die Retest-Reliabilität, sind im quantitativen Bereich ohnehin nur als Notbehelf zu betrachten; siehe die kritischen Einwürfe Mayrings 1996, 116; 2000, 110.),

- dass abhängig von der Bedeutsamkeit des Forschungsfelds in Studien zunächst *geringe Niveaus* der Kriterienbeachtung mangels besser geeigneter oder ökonomischerer Alternativen im Zuge einer Güterabwägung in Kauf genommen werden (eine im quantitativen Bereich nicht unübliche Praxis, z.B. die indirekte Erfassung objektiver Sachverhalte über Indikatoren der subjektiven Einschätzung) oder
- dass auch *ungeprüfte Instrumente* eingesetzt werden, ein Vorgehen, das in der quantitativen Methodik durchaus ebenfalls praktiziert wird, etwa bei einfachen, wenige Items umfassenden ad-hoc-Skalen mit einer überzeugenden Face-Validity (z.B. Kunter u.a. 2002).

Nicht im Sinne der Propagierung neuer Operationalisierungen ist es allerdings, Gütekriterien „methodenfreundlicher“ umzudefinieren: Konsensuell eingeführte Termini, wie z.B. „Gültigkeit“, semantisch anders zu besetzen, dürfte vor allem Verwirrung stiften (vgl. Lamnek 2005, 143, 148ff).

Es macht einen erheblichen Unterschied, ob Gütekriterien *grundsätzlich* abgelehnt und damit für irrelevant erklärt werden oder ob man sie zwar anerkennt, aber Risiken ihrer Verletzung in Ermangelung adäquaterer Erhebungsverfahren in *Einzelfällen* einzugehen bereit ist. Kriterien, deren Beachtung nicht geprüft wird bzw. werden kann, sind dadurch entsprechend gewürdigt, dass potentielle Verstöße gegen sie offengelegt bzw. Informationen zur Beurteilung der Störanfälligkeit des Vorgehens gegeben werden. So wird etwa auch in der quantitativen Schulforschung im Zusammenhang mit dem expliziten Einsatz „hoch-inferenter“, d. h. stark interpretationshaltiger und daher potentiell wenig objektiver Erhebungsinstrumente vorgegangen. Ein Beurteilungskriterium hingegen, das als Qualitätsstandard für Forschung zurückgewiesen wird, bedarf aus Sicht des Ablehnenden keiner weiteren Erwähnung.

Die pauschale Absage an Gütekriterien beruht wohl häufig auf der Verwechslung der einzelnen Kriterien an sich (z.B. Validität) mit ihren gängigen Operationalisierungsformen (z.B. Validierungsformen). Dieses Missverständnis wird bedauerlicherweise terminologisch nahegelegt. So sind etwa „Validierungsformen“, wie z.B. die Konstrukt- und Inhaltsvalidität oder die (öko)logische Validität (Kvale 1995), betitelt, als ob es sich um unterschied-

liche „Validitäten“ handeln würde (vgl. Saldern 1995, 353) und nicht um lediglich operationale Deutungen desselben Dachkonzepts.

Das meist umstrittene Gütekriterium dürfte die Objektivität sein. Kritiker des Objektivitätsgebots führen an, dass Forschung als menschliche Unternehmung immer von Subjekten durchgeführt wird und insofern zwangsläufig „subjektiv“ ist (vgl. Depaepe 1999, 190; Brezinka 1989, 140-142; Flick 2004, 20; Lamnek 2005, 172ff; Groeben 2006, 8). Mit Objektivität ist in diesem Zusammenhang jedoch nicht die „Absenz von Subjektverhaftetheit“ gemeint, sondern „Beurteilerunabhängigkeit“ im Sinne der Freiheit von individuumspezifischer Einfärbung, die nicht prinzipiell mit anderen Individuen geteilt wird, also im Kantschen Sinne „unabhängig von der Willkür des einzelnen“.

Das Wissenschaftskriterium der Objektivität – unmissverständlicher ausgedrückt, der Intersubjektivität - ist als Leitidee beachtenswert, wohl wissend, dass ihm in den Verhaltens- und Sozialwissenschaften selten absolut entsprechen werden kann. Deskriptive Theorien beanspruchen per se, reale Verhältnisse möglichst authentisch, also letztlich ‚wahrheitsgemäß‘ zu modellieren. „Beschreibende“ Aussagen, welche diesen Anspruch radikal aufgeben, mögen von ästhetisch-kontemplativem Wert sein (Huber 2001; vgl. Roeder 2006, 640). Orientierungspraktisch sind sie in der Regel sinnleer. Denn wozu dient ein Portrait, das nicht ambitioniert ist, den Portraitierten so gut als möglich abzubilden? Der radikal-konstruktivistische bzw. postmoderne Zweifel an einer objektiven, bewusstseinsunabhängigen Außenwelt steht dem nicht entgegen. Selbst wenn die gemeinsam erfahrbare Realität nur eine globale Hilfskonstruktion sein sollte, so zeigt bereits das Alltagserleben, dass diese Konstruktion nicht nur möglich, sondern als Verständigungsbasis sozial geradezu unentbehrlich ist. Insofern zielt der Verzicht auf „abbildtheoretische Wahrheitsauffassungen“ (Krüger 2000, 337) am Hauptzweck empirischen Vorgehens vorbei, das zwangsläufig auf diese Hilfskonstruktion beschränkt bleiben muss (vgl. Hammersley 1992, 50-52; Seale 1999, 23, 157). Auch „subjektive Wirklichkeiten“ als Forschungsgegenstände bilden hier keine Ausnahme. Soll die Weltkonstruktion einer Person (z. B. Meyer 2009, 310f) empirisch ermittelt werden, so wird ihre Weltsicht zu einem Teil der (kognitiven) Realität, die es gilt, möglichst zutreffend, also auch intersubjektiv zu beschreiben. Das Einfangen solcher subjektiver Konstruktionen markiert durchaus keinen separierbaren Gegenstandsbereich qualitativer Forschung, da dies bereits mit der quantitativen Erfassung von Fähigkeitsselbstkonzepten, Selbstwirksamkeitsüberzeugungen, Kausalattributionen oder Schulleistungsselbsteinschätzungen beginnt.

Es käme einem schroffen Bruch mit dem Objektivitätspostulat gleich, bewusste Parteinahme in der deskriptiven Forschung zuzulassen. Ein Beispiel: Zu dieser Frage erschienen im Rahmen der Koedukationsdebatte in der Zeitschrift für Pädagogik 1998 (243ff) Entgegnungen auf einen Artikel von Drerup (1997). Er hatte angemerkt, dass ein Teil der sich als feministisch verstehenden Schulpädagogik Züge von „Bewegungs- und Verständigungsliteratur“ aufweise, die stärker am politisch-praktischen Orientierungsbedürfnis ausgerichtet sei als am unvoreingenommenen Erkenntnisanspruch. (Ähnliches hatte Breitenbach bereits 1994 in einem nicht minder beachteten Periodikum formuliert, ohne eine größere Kontroverse damit zu entfachen.)

Auf Drerups Aufsatz erwiderten etliche Autorinnen, sie würden sich gegen ein Auseinanderdividieren von wissenschaftlichem und politischem Anspruch wehren, da gerade ihr gesellschaftliches Engagement und ihr Beitrag zur Verbesserung sozialer Verhältnisse Motor für ihre Forschung sei. Davon abgesehen, dass Parteilichkeit auch innerhalb der - teils qualitativ ausgerichteten - schulpädagogischen Geschlechterforschung nicht konsensfähig ist (Glumpler 1994, 20; vgl. Baumert 1992, 86), lassen sich beide Standpunkte zumindest in gewissem Umfang miteinander versöhnen. Objektivität fordert keine persönliche politische Abstinenz. Dies wäre ohnehin illusorisch (Popper 1969, 281). Engagement für die Sache und für gesellschaftliche Veränderung darf selbstverständlich im Entdeckungszusammenhang zum Ausdruck kommen, etwa auch in der Wahl der Forschungsthemen. Lediglich im Begründungszusammenhang von Forschung, wo es um die Prüfung von Annahmen geht, sollten persönliche Einstellungen nicht die kritische Sicht auf die Wirklichkeit vernebeln. So sollte beispielsweise unterbleiben, dass zeitgeist-konforme Meinungen ungeprüft als wahr vorausgesetzt werden, dass nur nach Bestätigung der eigenen Vormeinung gesucht wird, dass Gegenargumente - etwa durch selektive Literaturrezeption - verschwiegen werden (vgl. Baumert 1992, 86) und Mythenbildung durch Fahndung nach „selbstversteckten Ostereiern“ (Ulich u.a. 1981, 78) gefördert wird. Dies ist in der politischen Auseinandersetzung üblich und gilt dort als nicht unredlich. In dem Maße jedoch wie ein solches Agieren im wissenschaftlichen Diskurs praktiziert und damit die Trennung zwischen Politik und Wissenschaft aufgehoben würde, entzöge sich Wissenschaft die eigene Legitimationsbasis als eine für fundierte Erkenntnis zuständige Instanz des öffentlichen Lebens (vgl. Wottawa 2006, 162). Eine Vermischung von Politik und Wissenschaft „entwertet nicht nur die Forschung, sondern auch die Praxis“ (Breitenbach 1994, 184).

4. Apologie der universellen Position zu Forschungsstandards

Es ist davon auszugehen, dass der Korpus der Wissenschaftskriterien im Sinne idealtypischer Leitlinien, wie er von der quantitativen Methodik akzeptiert wird, historisch nicht zufällig zustande gekommen ist, etwa durch prinzipiell austauschbare Einflusskonstellationen oder durch beliebige paradigmatische Modeströmungen. Vielmehr ist dieser Kriterien-Katalog in den Augen der Mehrheit der scientific community offenbar so gut fundiert, dass er methodologie-übergreifend Anspruch auf allgemeine Anerkennung erheben kann (z. B. Brühl/Buch 2006; Lüders 2000, 634; Groeben 2006, 8). Einige dieser Standards werden gedeckt durch die Deutsche Gesellschaft für Erziehungswissenschaft (1997), die Kommission „Selbstkontrolle in der Wissenschaft“ der Deutschen Forschungsgemeinschaft, durch eine Resolution des Deutschen Hochschulverbandes (2000, 292) sowie international etwa durch den Bericht des amerikanischen National Research Council 2002 zum Stand der erziehungswissenschaftlichen Forschung in den USA (Feuer u.a. 2005, 31). Dieser Katalog wird also auch interdisziplinär anerkannt. Nicht nur die quantitative Forschung der Erziehungswissenschaft und ihrer Bezugsdisziplinen Psychologie und Soziologie orientiert sich an ihm, sondern alle Real- und Humanwissenschaften, z.B. die Meinungs- und Evaluationsforschung, die Medizin oder die Wirtschaftswissenschaften. (Die Naturwissenschaften als vermeintliches Vorbild quantitativer Ansätze in den Verhaltens- und Sozialwissenschaften [so z. B. Elliott et al. 1999, 217] teilen zwar dieselben empirischen Standards, bilden aber ansonsten weit weniger deren Ursprung als gemeinhin unterstellt wird; siehe Ludwig 2001, 83-85, 2003.)

Die Konsenswürdigkeit dieser Forschungskriterien lässt sich unschwer durch Zurückweisung ihrer Reversion demonstrieren. Denn jede gegenteilige Forderung wäre geradezu absurd: Niemand würde für wissenschaftliches Arbeiten ernsthaft etwa ein Höchstmaß an unsystematischer Subjektivität, Irrationalität oder Unkontrollierbarkeit postulieren.

Auch qualitative begründungskontextuelle Forschung kann sich insofern schwerlich diesen Standards entziehen, als sie – ob deren Vertreter dies begrüßen oder nicht - ohnehin an ihnen gemessen und nach ihnen beurteilt wird (Brühl/Buch 2006, 1; vgl. Schwartz-Shea 2009, 39). Daher sollten diese Kriterien innerhalb der qualitativen Methodologie angemessen diskutiert werden (Gehlert 2000, 72; Prein/Erzberger 2000, 346). Dies geschah bislang erst ansatzweise (z.B. Flick 2000; 2004, 98; Bohnsack 2005, 76f) und auch nicht immer, ohne dass die Argumentation hinter das schon andernorts erreichte wissenschaftstheoretische Niveau zurückfällt (z.B. Denzin 2000, 147).

5. Implikationen der Absenz eines Standard-Konsenses

Sowohl quantitative als auch qualitative Studien weisen eine erhebliche Spannbreite hinsichtlich ihrer methodischen Güte und damit der Aussagekraft ihrer Befunde auf (Lüders 2000, 635; Reichertz 2000, 13f, 17). Eisner reduziert diese Spannbreite beim qualitativen Paradigma auf zwei Qualitätsniveaus: den „*scientific*“ und den „*artistic*“ approach (1981). Während jedoch die gemeinsame Anerkennung von Standards im quantitativen Bereich eine Beurteilung der Resultate nach ihrer Tragfähigkeit und damit die Bewertung der Studien als mehr oder weniger „stark“ bzw. „schwach“ erlaubt, ist dies im qualitativen Bereich aufgrund der dort wenig geführten Binnendiskussion zu Qualitätsstandards nicht in derselben Weise möglich. Daher stellt „eines der grundlegenden Probleme der qualitativen Forschung (...) die Bewertung ihrer Ergebnisse dar“ (Knoblauch 2000, 628; siehe auch Flick 2002, 317). Forschungsergebnisse, die überhaupt nicht bewertet werden könnten, wären nahezu unbrauchbar. Denn erst durch eine solche Beurteilbarkeit lässt sich die einfache, aber entscheidende Frage beantworten, inwieweit den Erträgen zu trauen ist. Eine brückenbildende Formulierung Kurt Weills zur Überwindung der Spaltung der Tonkunst in sogenannte *unterhaltende* und *ernste* Musik könnte analog auf den vorliegenden Kontext übertragen werden: Es ist wichtiger, zwischen brauchbarer und ertragsarmer Forschung zu unterscheiden als zwischen qualitativer und quantitativer.

Der Mangel an geteilten Qualitätsstandards mag dazu beitragen, dass der qualitative Forschungsbereich anscheinend in Verkennung seines Anforderungsniveaus mitunter eine gesteigerte Attraktivität auf methodisch Unbedarfte ausübt (Hopf/Müller 1995, 66; Elliott et al. 1999, 218), die sich der Illusion hingeben, „qualitativ zu forschen sei leichter“ (Oswald 1997, 71). „Leider verbirgt sich hinter der Maske qualitativer Forschung häufig nur blanker Dilettantismus“ (Wellenreuther 2000, 14; ebenso Reichertz 2000, 14). „Manche eher schlichte qualitative Projekte machen dies deutlich – sei es dadurch, daß hohles Theoriebildungspathos mit naiver Deskription verbunden wird, oder dadurch, daß die entsprechenden Forschungsberichte eher Dokumente der Vorurteilsproduktion als der Dateninterpretation sind“ (Hopf/Müller 1995, 66f). Daraus sollten allerdings keine vorschnellen „Schlüsse zu den Entwicklungschancen“ und zur Komplexität qualitativer Forschung gezogen werden (ebd.).

6. Anforderungen an qualitative Forschung

Die Diskussion um die Etablierung einer Kultur von Standards für die Anwendung qualitativer Verfahren im Begründungszusammenhang kann durch die Reflexion traditioneller Forschungspostulate bereichert werden. Die folgende exemplarische *Postulate-Sammlung* wurde aus Defiziten abgeleitet, wie sie typischerweise in schwachen qualitativen Untersuchungen zu finden sind. Diese Forderungen sind also keineswegs neu. Mit ihrer Auflistung soll unterstrichen werden, dass die Forschungspraxis nach wie vor auf die Notwendigkeit hinweist, derartige Aspekte in einen Katalog von Qualitätsangaben einzubinden. Weitere Punkte finden sich etwa bei Krüger (2000, 337) und Seale (1999). (Analoge Auflistungen lassen sich natürlich ebenso für quantitative Arbeiten erstellen; z.B. Rost 2005; Stelzl 2005.)

(1) Auch qualitative Studien sollten ihr zentrales Erkenntnisziel bzw. ihren Erkenntnisanspruch explizit offenlegen. Bisweilen lassen Forschungsberichte weder am methodischen Vorgehen noch an der Ergebnisdarstellung oder – zusammenfassung klar erkennen, worum es bei der Studie geht: Dient die Untersuchung etwa dazu, neue Fragen aufzuwerfen bzw. (kausale) Annahmen aufzustellen oder Hypothesen zu prüfen bzw. Fragen zu beantworten (z. B. Waider 2005)? Ist die Untersuchung im Begründungszusammenhang anzusiedeln, sollte dargelegt werden, welcher Anspruch an die Belastbarkeit der Befunde gestellt wird bzw. werden kann: Handelt es sich um possibilistische Erkenntnisse, mit denen aufgezeigt werden soll, dass z. B. bestimmte Lernresultate unter bestimmten Lernvoraussetzungen auftreten *können*, um damit harte „Unmöglichkeitsfalsifikationen“ vorzunehmen (vgl. Ludwig/Ludwig 2006, 194), oder um den untermauerten Beleg regelhafter Zusammenhänge zwischen Lernvoraussetzungen und -resultaten? Stellen die Einzelfalldokumentationen (z.B. aus Unterrichtsprotokollen) eine Beispielsammlung von nur möglichen oder typischen Phänomenen des Untersuchungsgegenstands dar? Oder wird der Anspruch erhoben, in den erfassten Kategorien das Ereignisfeld erschöpfend dargestellt zu haben? Wird beansprucht, beispielsweise aus einer dokumentierten Kontingenz von Lehreraktion und Schülerreaktion eine plausible kausale Vermutung ableiten zu können oder einen stabilen Beleg für einen Kausalzusammenhang erbracht zu haben (z. B. Budde u.a. 2008, 248-255)?

(2) Die Auswertung qualitativer Daten im Begründungszusammenhang sollte, wenn nicht hypothesengeleitet, so doch stringent forschungsfragen-gerichtet erfolgen, um nach Mayring „nicht nur irgendwie interessantes Material (zu) präsentieren“ (1999, 296; vgl. Brüsemeister 2000, 227). Sehr allgemeine Fragestellungen, etwa Konkretisierungen für sonst nur abstrakt formulierte

Positionen aufzuspüren, z.B. „wie die Herstellung von Geschlecht in Kommunikationsprozessen funktioniert“ (Güting 2004), verleiten dazu, eine schwer handhabbare illustrative Beispielsammlung als Ergebnis zu offerieren.

(3) Qualitative Forschung hat sich an elementaren Mindestanforderungen zu orientieren, die generell an empirische Forschung gleich welcher Couleur gestellt werden. So manches „leitfadengeführte Experteninterview“ entpuppt sich bei näherer Betrachtung als quasi-journalistische Informationsbeschaffung, die wenig mit empirischem Vorgehen im Allgemeinen zu tun hat und daher auch keine qualitative Befragung (im Sinne eines wissenschaftlichen Erhebungsinstruments) im Besonderen darstellt, weil keine Primärdaten erfasst werden (vgl. Lüders 2000, 638). Empirische Forschung leitet ihre Befunde *unmittelbar* stringent aus einem Set erfahrener, registrierter und dokumentierter Einzelinformationen (Daten, Material) ab. Solche unmittelbaren Erfahrungen sind *Primärerfahrungen* mit dem Untersuchungsgegenstand. Ob die erfassten Daten Primärcharakter haben, hängt von der Fragestellung ab. Werden Experten befragt, erhält man in der Regel Primärdaten zur *Meinung* dieser Experten zu einem Sachverhalt, jedoch nicht zum *Sachverhalt* selbst. Ist dabei die (Detailinformation aggregierende) Expertenmeinung der Forschungsgegenstand, so handelt es sich um empirische Forschung. Geht es hingegen um die Prüfung einer Aussage zu einem pädagogischen Sachverhalt, so stellt die Befragung von pädagogischen Experten als Autoritäten kein empirisches Vorgehen dar. Aus diesem Grund kann das systematische Durchforsten einer Schrift von Johann H. Pestalozzi oder Maria Montessori eine quantitative oder qualitative Inhalts- oder Dokumentenanalyse und damit empirisches Forschen sein, soweit z.B. das Denken einer Person oder einer Epoche Forschungsgegenstand sind. Geht es hingegen etwa um die Prüfung von pädagogischen Theorien, mag der Standpunkt eines prominenten Fachvertreters als Gewährspersonen von *theoretischem* Interesse sein. Die Erkundung eines solchen Standpunkts stellt jedoch kein *empirisches* Prüfen dar, weil die Zustimmung etwa eines Schleiermachers zu einer These diese *erfahrungswissenschaftlich* nicht bekräftigt.

(4) Qualitatives Material sollte nicht mit selektivem Blick auf hypothesenbestätigende Fundstellen nur affirmativ ausgewertet werden. Es muss auch fallibilistisch gearbeitet und gezielt nach „Gegenevidenzen gefahndet“ werden (Prein/Erzberger 2000, 349; vgl. Kelle/Kluge 2006; Seale 1999, 73ff; Bohnsack 2008, 30; Schwartz-Shea 2009, 35). Beim empirischen Arbeiten ist die Ambiguitätstoleranz aufzubringen, dass – nach Thomas Henry Huxley (1870) – „wunderschöne Hypothesen von häßlichen Tatsachen ermordet werden“ können. Eine Maxime des Kritischen Rationalismus' ruft Forschende

auf, die potentiellen Prüfungsausgänge zu benennen, bei deren Auftreten sie bereit wären, die Hypothesen als widerlegt zu betrachten. Je weiter dieses Fenster der Falsifikationsmöglichkeiten geöffnet wird, umso tiefer ist das qualitative „Signifikanzniveau“ angesetzt bzw. desto stärker ist eine Entscheidung zugunsten der Hypothese gegen den „ α -Fehler“ gesichert (Cohen/Manion/Morrison 2007, 116f, 197, 221).

(5) Empirische Resultate sollten eine stringente Anbindung an das Datenmaterial aufweisen. Schlussfolgerungen stützen sich plausibel auf diese Daten. Allzu freie „tiefenhermeneutische“ Interpretationen oder Reflexionen über das Forschungsthema in lediglich grober Anlehnung an das erhobene Informationsmaterial, welche die verbalen Daten zur Aktivierung eigener Kreativität nutzen, mögen zu innovativem erkenntniserweiterndem Brainstorming führen. Mit *empirischer* (erfahrungsbezogener und -kontrollierter) Erkenntnisgewinnung hat ein derartiges Vorgehen oft nur noch wenig gemein (vgl. Rost 2005, 61; Eco 1987, 46ff; Soeffner 1989, 59). (Pointiert „autonome“ Interpretationen liefern bisweilen durchaus valide Auskünfte, allerdings weniger über den Interpretierten als vielmehr den Interpreten.) Dazu mögen Textinterpretationen gehören, welche sich damit begnügen, „stilisierte empirische Einblicke“ zu ermöglichen, bei denen „signifikante Eindrücke“, „nicht einfach über die Angabe einer handvoll Transkriptsequenzen belegt werden können“ (Mecheril 2003, 23, 45, 47). Zur selbstkritischen Kontrolle kann die Frage dienen: Sind die empirisch gewonnenen Ergebnisse abgesicherter als es freie theoretische Überlegungen wären, die durch bloßes Reflektieren entstehen (Prein/Erzberger 2000, 350)?

(6) Es ist erforderlich, Informationen zu liefern, welche den Grad der Gesicherheit der Resultate einschätzen lässt. Dazu gehört etwa die Angabe alternativer Schlussfolgerungen, welche die Zuverlässigkeit der Befunde bzw. deren „qualitative“ Irrtumswahrscheinlichkeit verdeutlichen. Den (knappen) veröffentlichten Berichtlegungen ist es oft nicht mehr anzusehen, inwieweit die Daten stringent zu den Interpretationen geführt haben. Deskriptive Resultate und Interpretationen sind getrennt darzustellen und deutlich als solche auszuweisen (vgl. Seale 1999, 191). Eine lediglich allgemeine (projektunspezifische) Einführung in die angewendete Methode, wie z.B. die ethnographische Analyse, reicht nicht aus, um die Stichhaltigkeit einzelner Deutungen prüfen zu können (Seale 1999).

(7) Nach Oswald ist bei qualitativer Forschung der „Ertrag oft ungewiß“, was teilweise auf den „unüberlegten Einsatz“ zurückzuführen ist (1997, 71). Auch qualitative Forschung sollte explizit resultate-fokussiert vorgehen (Krüger 2000, 337). Im Auswertungsprozess werden die Ausgangsinformationen

stringent auf ein Forschungsziel hin reduziert. Am Ende des Forschungsprozesses sollten relativ knapp gefasste spezifische Hauptergebnisse vorliegen: die Beantwortung oder Bearbeitung einer Hypothese, einer konkreten Fragestellung oder - bei hypothesengenerierender Forschung - empirisch begründbare Annahmen. Wenn Zusammenfassungen nahezu inhaltsleer abstrakt gehalten sind oder die „Befunde“ aus einer Fülle von (in den Einzelfallbeschreibungen verstreuten zusammenhangs-, theorie- und verallgemeinerungsarmen) Detailinformationen erscheinen, sind Rezipienten dazu verurteilt, einen Forschungsbericht von mehreren hundert Seiten zu studieren, nur um sich einen „ersten“ Überblick über eine Studie zu verschaffen. Wichtige Resultate, „verborgen“ in einem weitläufigen Textschungel von ausführlichen Einzelfalldarlegungen, sind fast ebenso unzugänglich wie noch unerforschte Felder. Forschungsergebnisse müssen kommunizierbar gehalten sein, d.h. auch einem „Sparsamkeitsgebot“ in der publizistischen Aufbereitung folgen (Deutsche Gesellschaft für Erziehungswissenschaft 2006, 34). Aussagefähige Bündelungen und damit Verdichtung und Komprimierung sind „unvermeidlich, wenn man zu sozialen Aussagen kommen will und nicht bei der Beschreibung der unendlichen Zahl individueller Lebensführungen“ (Preuss-Lausitz 1997, 434) oder narrativen Ereignisdokumentationen stehen bleiben will (z.B. Güting 2004; Krammling-Jöhrens 1997; Bannach 2002). Ausreichend dokumentierten „anekdotischen“ Einzelfällen wird durchaus auch in der quantitativen Forschung ein gewisser Aussagewert zugebilligt (vgl. Rosenthal 1976, 328; Uhl 1996, 21). Verwendbare Forschungsergebnisse sind orientierende, möglichst treffende „verkleinerte“, theoretisch eingedampfte Abbildungen einer meist weit komplexeren Realität. Denn „wer, so die provokativ überspitzte Frage, braucht eine Landkarte von Argentinien in der Größe von Argentinien?“ (Schnabel 1999, 336). Der Versuch einer Totaleruierung des gewählten Wirklichkeitsausschnitts, seine „holistische Erfassung“, scheitert meist, wenn nicht an der Akribie und Unermüdlichkeit der Forschenden, so doch spätestens an der überstrapazierten Geduld ihrer Rezipienten (vgl. Ludwig/Ludwig 2006, 185).

(8) Auch qualitative Studien sollten sich explizit zur Generalisierbarkeit ihrer zentralen Ergebnisse äußern (Flick 2000, 259f; 2002, 336ff; 2004, 98; Schwartz-Shea 2009, 38; Prein/Erzberger 2000, 346; Bohnsack 2005, 76f; Mayring 1999, 297; Lamnek 2005, 180ff; Seale 1999, 106-118). Dabei ist zu berücksichtigen, dass sich Einzelfallbetrachtung und nomothetische Orientierung keinesfalls gegenseitig ausschließen. Auch an einem einzigen Ereignis oder wenigen Fällen kann Grundsätzliches, Regelmäßiges oder Typisches studiert werden, so dass die gewonnenen Erkenntnisse gegebenenfalls auf nicht untersuchte strukturverwandte Fälle in gewissem Umfang

übertragbar sind (Kelle/Kluge 2006; vgl. Bastian u.a. 2000, 121; Horstkemper/Tillmann 2003, 45, 47; Kirchhöfer 1997, 208; Lamnek 1995, 221ff; Leutner 1999, 331; Kardorff 1995, 5). Denn weder qualitative noch quantitative Untersuchungen, die uneingeschränkt idiographisch orientiert sind, lohnen ihren Aufwand. Auch Einzelfallforschung verfolgt - soweit sie über das Selbstzweckdienliche hinausgeht - zumindest implizit das Interesse, über die konkret untersuchten Fälle hinaus, verallgemeinerbare, „nomothetische“ Kenntnisse zu erlangen.

7. Fazit

In der quantitativen Forschung ist eine lebendige Tradition gegenseitiger kritischer Würdigung hinsichtlich der Entwicklung und Anwendung von Methoden etabliert (z.B. Rost 2005; Stelzl 2005; Roeder 2006, 642; Cohen/Manion/Morrison 2000, 5). Dies gilt bislang nicht in gleichem Umfang für den qualitativen Forschungsbereich. Dort scheint noch weitgehend eine Neigung zur einvernehmlichen, aber bezugsfreien Koexistenz und zur unhinterfragenden Duldsamkeit gegenüber methodischen Auffassungen vorzuherrschen (vgl. Groeben 2006, 5f). Dies mag ihre Ursache auch darin haben, dass sich innerhalb der qualitativen Methodologie bisher kaum eine Einigkeit über Standards abzeichnet und damit keine einheitlichen Qualitätsmaßstäbe zur Verfügung stehen, von denen aus Bewertungen vorgenommen werden könnten. Die in den letzten Jahren zahlreich vorgelegten einführenden Übersichtswerke zu qualitativen Methoden sind mit (selbst)kritischen Auseinandersetzungen, etwa über die Reichweite (Möglichkeiten und Einschränkungen im Einsatz) einzelner Methoden, eher zurückhaltend (z.B. Breidenstein u.a. 2002). Eine solche Laissez-Faire-Haltung ist der Forschungsk Kooperation wenig zuträglich. Die Absenz von gemeinsamen methodologischen Bezugspunkten mag ein Grund dafür sein, dass einzelne qualitative Forschungsbemühungen meist isoliert bleiben, kaum Anschlussprojekte entstehen oder institutsübergreifende Forschungslinien und –netzwerke aufgebaut werden (Krüger 2000, 337f; vgl. Helsper/Herwartz-Emden/Terhart 2001, 253, 264f). Aussichtsreich erscheint daher eine stärkere Förderung der Kultur gegenseitiger forschungsmethodischer und –methodologischer Kritik. Denn „wer Kritik nicht will, will auch eigentlich (im wissenschaftlichen Sinn) gar nichts wissen“ (Lochner 1963, 9; vgl. Seale 1999, 30). Eine solche Kultur hätte eine selbstprotektive Funktion: Sie kann durch klare Grenzziehung verhindern, dass das Etikett „qualitativ“ als Immunisierungsstrategie zur Legitimation methodisch fragwürdiger Arbeiten verfügbar ist und dass nach Lüders „unter

dem Label qualitative Sozialforschung methodologisch noch immer scheinbar nahezu alles möglich ist“ (2000, 635) bis hin zu „esoterischen Kunstlehren“ (Groeben 2006, 1).

Erziehungswissenschaftliche Forschung im Entdeckungszusammenhang erzeugt Annahmen, Vermutungen und eröffnet damit Blickfelder, Denkmöglichkeiten und potentielle Sichtweisen. Sie zeigt, wie die pädagogische Welt *vielleicht aussieht*. Der Erkenntnisanspruch im Begründungszusammenhang hingegen ist darauf gerichtet herauszuschälen, wie die Welt *tatsächlich ist*. Anwendungsbezogene empirisch-pädagogische Forschung mit diesem Anspruch unterliegt per se Qualitätsstandards, da sie letztlich auf die Generierung möglichst gesicherten verbindlichen Wissens zielt und damit wiederum evidenz-basierte Erziehung und Bildung (Cohen/Manion/Morrison 2007) ermöglichen will.

Derzeit sind deutlich Bestrebungen zu erleben, auch solche Bereiche des öffentlichen Lebens durchökonomisieren zu wollen, welche marktwirtschaftlichen Prinzipien nicht per se gehorchen (vgl. Ludwig/Ludwig 2006, 185), wie etwa die Forschungs- und Hochschullandschaft. Jedoch auch aus der kritisch-distanzierter Sicht auf diese neoliberale Tendenz wird man der Maxime, Forschung sei gehalten, wirtschaftlich mit Zeit und Ressourcen umzugehen, die Zustimmung nicht ganz verweigern können. D.h. zwischen dem investierten Forschungsaufwand und dem Grad der Relevanz und Belastbarkeit des erwartbaren Forschungsertrags sollte keine eklatante Schiefelage entstehen (Oswald 1997, 71; Lüders 2000, 636; Helsper/Herwartz-Emden/Terhart 2001, 264). Möglicherweise wäre ein Teil des immensen Aufwands, den einige qualitative Verfahren allein zur Hypothesengenerierung erfordern, bisweilen gewinnbringender in eine nachfolgende Hypothesenprüfung investiert. Der Eindruck einer ungerechtfertigten „Kosten-Nutzen-Relation“ (Oswald 1997, 71) stellt sich dort ein, wo nicht recht einleuchten will, warum die mit hohem methodisch-planvollen Einsatz produzierten Hypothesen grundsätzlich realitätsangemessener sein sollen als Thesen, die feldsensiblen, kreativen Spontan-Einfällen mittels der „Drei-B-Methode“ zu verdanken sind (in *bed*, in *bathroom*, on *bike*; Bortz/Döring 2006, 353). Bisweilen scheint der *Weg* der methodologischen Selbstreferentialität und – reflexion verbunden mit einer Freude an der Ästhetik der eigenen Interpretationskunst das eigentliche *Ziel* von so manchem Methodenenthusiasten zu sein. (Ein Analogon lässt sich dazu durchaus in quantitativen Untersuchungen finden, etwa wenn sich die Vermutung aufdrängt, bereits eine gründlichere a-priori-Sichtprüfung hätte die Unangemessenheit von Hypothesen oder Testitems auch ohne Rückgriff auf komplexe statistische Verfahren zu Tage bringen können.)

Forschungsansätze etablieren sich vermutlich letztlich weniger durch programmatisch-methodologische Darstellungen, als vielmehr durch „überzeugende Erkenntnisgewinne, die durch die jeweiligen Verfahren erzielt werden“ (Hopf/Müller 1995, 67). Nach einer gewissen Entwicklungs- und Testphase müssen es sich Forschungsmethoden gefallen lassen, nach den mit ihnen produzierten Forschungsergebnissen beurteilt zu werden. Für die resümierende Einschätzung der Tragfähigkeit einzelner qualitativer Forschungsverfahren wären kritisch-bilanzierende Kompendien zur Best-Practice-Anwendung der Verfahren mittels exemplarischer, ausführlich beschriebener Studien, die zu gehaltvollen, praxisrelevanten bildungswissenschaftlichen Erträgen führten, derzeit hilfreicher als weitere rein method(olog)ische Sammelbände (vgl. Helsper/Herwartz-Emden/Terhart 2001, 259, 266). Die inhaltliche Auswahl von Studien für solche Übersichtswerke setzt wiederum ein Übereinkommen zu Qualitätsstandards voraus.

Die Diskussion über Qualitätsstandards und Bewertung qualitativer Forschung ist bereits angestoßen (z.B. Elliott et al. 1999, 219ff; 2000; Seale 1999; Schwartz-Shea 2009). Vorgelegte Taxonomien zu Qualitätskriterien beinhalten jedoch teils noch sehr vage abstrakt formulierte Elemente, etwa die nicht näher erläuterte Forderung nach „geeigneter“ Ergebnisdarstellung, einer „adäquaten“ Auswahl von Feldzugängen oder nach dem „Grad der Angemessenheit an die empirische Realität“ (Elliott et al. 1999; Seale 1999, 189ff), nach „Wahrheit“, „Glaubwürdigkeit“ oder „Erkennbarkeit“ (z.B. Biller 1988, 45ff). Teils enthalten sie triviale Postulate wie „Plausibilität“, „Authentizität“ und „Gegenstandsbezogenheit“ (z. B. Lüders 2000, 633f). Teils werden eher Fragen gestellt als Antworten gegeben (z. B. Helsper/Herwartz-Emden/Terhart 2001, 261-265). Zusammenfassend kann festgehalten werden: Es wäre begrüßenswert, wenn diese Diskussion innerhalb der qualitativen Richtung nachdrücklicher geführt und idealiter in einer gemeinsam oder zumindest mehrheitlich getragenen Grundposition münden würde. Dies würde nach Meinung Elliotts und Kollegen in erheblichem Umfang zur Legitimation, und einer breiteren Anerkennung und Akzeptanz von qualitativer Forschung beitragen (1999, 217, 225). Die Verständigung auf eine gemeinsame Linie würde wissenschaftstheoretisch ein tragfähiges Fundament setzen, auch, um den bis heute noch wenig konsensuell geklärten Begriff „qualitative Forschung“ fassen zu können (vgl. Ludwig 2004). Ein unklarer Begriff an prominenter Stelle ist kommunikationshinderlich. Ohne ein solches Fundament droht der Terminus „qualitativ“ zur methodischen Restkategorie zu mutieren.

Selbst wenn die Verständigung auf eine gemeinsame Position ausbleibt, wäre bereits dann ein erheblicher Fortschritt erzielt, soweit sich die einzelnen

qualitativen Schulen klar und differenziert dazu äußern würden, welche Standards sie für verbindlich für die Anwendung ihrer Methodeninstrumentariums erklären. Dadurch würden die Rezipienten ihrer Studien leichter erkennen können, ob zu erwarten ist, dass sich die im Begründungskontext von Forschung angewendeten qualitativen Verfahren im Rahmen der gängigen, allgemein akzeptierten Wissenschaftsauffassung bewegen und daher deren Befunde theoretisch und empirisch anschlussfähig sind. „Nur wenn die Standards wissenschaftlicher Güteprüfung in der qualitativen Forschung fest etabliert auch weiter ausdifferenziert werden“, hat das qualitative Methodologieprogramm „eine Chance, auf dem Markt zu bleiben und dort zu bestehen“ (Reichertz 2000, 18; vgl. Ludwig 2007; 2009).

Anmerkung: Die in den vorausgehenden Ausführungen wörtlich bzw. sinngemäß wiedergegebenen Einschätzungen zur Rezeption und Wahrnehmung qualitativer Forschung, zur Anwendungspraxis qualitativer Methoden und zum Diskussionsstand von Qualitätskriterien innerhalb des qualitativen Paradigmas stammen fast ausschließlich von selbst qualitativ orientierten Autorinnen und Autoren.

LITERATUR

- Abel J./Möller R./Treumann K. (1998): Einführung in die empirische Pädagogik. Stuttgart
- Bamme A. (1986): Anything goes - science everywhere? Konturen von Wissenschaft heute. München
- Banister P./Burman E./Parker I./Taylor M./Tindall C. (1994): Qualitative methods in psychology. A research guide. Buckingham
- Bannach M. (2002): Selbstbestimmtes Lernen. Baltmannsweiler
- Bastian J./Combe A./Gudjons H./Herzmann P./Rabenstein K. (2000): Profile in der Oberstufe. Hamburg
- Baumann U. (1998): Kausalität und qualitative empirische Sozialforschung. Das Verstehen im Dienst der Ursache-Wirkungs-Forschung und die Intentionalität. Münster
- Baumert J. (1992): Koedukation oder Geschlechtertrennung? Zeitschrift für Pädagogik, 38. Jg., Nr. 1, 83-110
- Beutelspacher A. (2002): Auch der Zufall gehorcht mathematischen Gesetzen. Forschung & Lehre 2, 65-67
- Biller K. (1988): Pädagogische Kasuistik. Baltmannsweiler
- Blömeke S. (2004): Empirische Befunde zur Wirksamkeit der Lehrerbildung. In: Blömeke S./Reinhold P./Tulodziecki G./Wildt J. (Hrsg.): Handbuch Lehrerbildung. Bad Heilbrunn, 59-91
- Bohnsack R. (2005): Standards nicht-standardisierter Forschung in den Erziehungs- und Sozialwissenschaften. Zeitschrift für Erziehungswissenschaft 8, Beiheft 4, 63-81

- Bohnsack R. (2008): Rekonstruktive Sozialforschung. Einführung in qualitative Methoden. Opladen
- Bortz J./Döring N. (2006): Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. Berlin
- Breidenstein G./Combe A./Helsper W./Stelmaszyk B. (Hrsg.) (2002): Forum qualitative Schulforschung 2. Interpretative Unterrichts- und Schulbegleitforschung. Opladen
- Breitenbach E. (1994): Geschlechtsspezifische Interaktion in der Schule. Eine Bestandsaufnahme der feministischen Schulforschung. Die deutsche Schule, 86. Jg., H. 2, 179-191
- Breuer F./Kölbl C./Reichertz J./Roth W. (o.J.): Debatte „Qualitätsstandards qualitativer Sozialforschung“. Moderation. Forum Qualitative Sozialforschung (on-line Journal) [www.qualitative-research.net/fqs-d/debate-1; Zugriff: 23.3.2007].
- Breuer F./Reichertz J. (2001): Wissenschaftskriterien: Eine Moderation. Forum Qualitative Sozialforschung (on-line) 2, H. 3
- Brezinka W. (1989/1): Aufklärung über Erziehungstheorien. Gesammelte Schriften Bd. 1. München
- Bridges D. (2008): Evidence-based Reform in Education: a response to Robert Slavin. European Educational Research Journal 7, No. 1, 129-133
- Brühl R./Buch S. (2006): Einheitliche Gütekriterien in der empirischen Forschung? Objektivität, Reliabilität und Validität in der Diskussion. ESCP-EAP - Europäische Wirtschaftshochschule Berlin. Working Paper No. 20. Dezember 2006 [Zugang: http://www.escp-eap.eu/uploads/media/Bruehl_Buch_Einheitliche_Guetekriterien_2006_04.pdf; Zugriff am 1.3.2010]
- Brüsemeister T. (2000): Qualitative Forschung. Ein Überblick. Wiesbaden
- Budde J./Scholand B./Faulstich-Wieland H. (2008): Geschlechtergerechtigkeit in der Schule. Weinheim
- Cohen L./Manion L./Morrison K. (2007): Research methods in education. London
- Combe A./Helsper W./Stelmaszyk B. (Hrsg.) (1999): Forum Qualitative Schulforschung 1. Weinheim
- Creswell J.W. (2003): Research design – qualitative, quantitative, and mixed methods approaches. Thousand Oaks
- Danner H. (2006): Methoden geisteswissenschaftlicher Pädagogik. München
- Denzin N. (1970): The research act. Englewood Cliffs, NJ
- Denzin N. (2000): Symbolischer Interaktionismus. In: Flick/Kardorff/Steinke (Hrsg.), 136-150
- Depaepe M. (1999): Experimentelle Pädagogik, Reformpädagogik und pädagogische Praxis. In: Oelkers J./Osterwalder F. (Hrsg.): Die neue Erziehung. Beiträge zur Internationalität der Reformpädagogik. Bern, 183-205
- Deutsche Gesellschaft für Erziehungswissenschaft (1997): Standards erziehungswissenschaftlicher Forschung. In: Friebertshäuser B./Prengel A. (Hrsg.): Handbuch Qualitative Forschungsmethoden in der Erziehungswissenschaft. Weinheim, 857-863
- Deutsche Gesellschaft für Erziehungswissenschaft: der Vorstand (2006): Anonymisierung von Daten in der qualitativen Forschung: Probleme und Empfehlungen. Erziehungswissenschaft 17, H. 32, 33-34
- Deutscher Hochschulverband (2000): Wissenschaftliches Fehlverhalten. Resolution des 50. Hochschulverbandstages vom 5. April 2000 in Berlin. Forschung & Lehre 6, 292
- Diekmann A. (2001): Empirische Sozialforschung. Reinbek

- Drerup H. (1997): Die neue Koedukationsdebatte zwischen Wissenschaftsanspruch und politisch-praktischem Orientierungsbedürfnis. *Zeitschrift für Pädagogik*, 43. Jg., Nr. 6, 853-875
- Eco U. (1987): Streit der Interpretationen. Konstanz
- Eisner E. (1981): A rejoinder. *Educational Researcher* 10, 25-26
- Elliott R./Fischer C./Rennie D. (1999): Evolving guidelines for publication of qualitative research studies in psychology and related fields. *British Journal of Clinical Psychology* 38, 215-229
- Esser H. (1987): Zum Verhältnis von qualitativen und quantitativen Methoden in der Sozialforschung, oder: Über den Nutzen methodologischer Regeln bei der Diskussion von Scheinkontroversen. In: Voges W. (Hrsg.): *Methoden der Biographie- und Lebenslaufforschung*. Opladen, 87-101
- Eye A. v. (1994): Zum Verhältnis zwischen qualitativen und quantitativen Methoden in der empirisch-pädagogischen Forschung. In: Olechowski R./Rollett B. (Hrsg.): *Theorie und Praxis*. Frankfurt a.M., 24-45
- Feuer M./Towne L./Shavelson R. (2005): Scientific Culture and Educational Research. *Zeitschrift für Erziehungswissenschaft* 8, Beiheft 4, 25-44
- Feyerabend P. (1979): *Erkenntnis für freie Menschen*. Frankfurt a.M.
- Feyerabend P. (1983): *Wider den Methodenzwang. Skizze einer anarchistischen Erkenntnistheorie*. Frankfurt a.M.
- Feyerabend P. (1995): *Zeitverschwendung*. Frankfurt a.M.
- Feyerabend P./Thomas, C. (Hrsg.) (1986): *Nutzniesser und Betroffene von Wissenschaften*. Zürich
- Finkbeiner C./Ludwig P.H./Wilden E./Knierim M. (2006): ADEQUA - Bericht über ein DFG-Forschungsprojekt zur Förderung von Lernstrategien im Englischunterricht. *Zeitschrift für Fremdsprachenforschung* 17, H. 2, 257-274
- Flick U. (2000): Design und Prozess qualitativer Forschung. In: Flick U./Kardorff E.v./Steinke J. (Hrsg.): *Qualitative Forschung*. Reinbek, 252-265
- Flick U. (2002): *Qualitative Sozialforschung. Eine Einführung*. Reinbek: Rowohlt
- Flick U. (2004): *Triangulation*. Wiesbaden
- Flick U./Kardorff E.v./Steinke J. (Hrsg.) (2000): *Qualitative Forschung*. Reinbek
- Frey H.P. (1983): *Stigma und Identität*. Weinheim
- Früh W. (1991): *Inhaltsanalyse*. München
- Gehlert S. (2000): Computerunterstützte Auswertung qualitativer Daten. Anforderungen, Realisierungsmöglichkeiten, Grenzen. In: Macha H. und Forschungsgruppe: *Erfolgreiche Frauen*. Frankfurt a.M., 72-98
- Glumpler E. (1994): Koedukation. Entwicklungen und Perspektiven. In: Glumpler E. (Hrsg.): *Koedukation*. Bad Heilbrunn, 9-30
- Groeben N. (2006). Gibt es Wege aus der selbstverschuldeten Irrelevanz des qualitativen Offstreams? *Forum Qualitative Sozialforschung (On-line Journal)* 7, No. 4
- Gütting D. (2004): Soziale Konstruktion von Geschlecht im Unterricht. *Ethnographische Analysen alltäglicher Inszenierungspraktiken*. Bad Heilbrunn
- Hammersley M. (1992): *What's wrong with ethnography?* London
- Hammersley M. (1996): The relationship between qualitative and quantitative research. In: Richardson J.T. (ed.): *Handbook of qualitative research methods for Psychology and the Social Sciences*. Leicester, 159-174
- Helfferich C. (2004): *Die Qualität qualitativer Daten*. Wiesbaden
- Helmke A./Schrader F.-W. (2006): Determinanten der Schulleistung. In: Rost D.H. (Hrsg.): *Handwörterbuch Pädagogische Psychologie*. Weinheim, 83-94
- Helsper W./Herwartz-Emden L./Terhart E. (2001): Qualität qualitativer Forschung in der Erziehungswissenschaft. *Zeitschrift für Pädagogik* 14, 251-269
- Hesse H. (1943): *Das Glasperlenspiel*. Zürich

- Hirschauer S. (2008): Die Empiriegeladenheit von Theorien und der Erfindungsreichtum der Praxis. In: Kalthoff H./Hirschauer S./Lindemann G. (Hrsg.): Theoretische Empirie: Zur Relevanz qualitativer Forschung. Frankfurt a.M., 165-187
- Hopf C. (1996): Hypothesenprüfung und qualitative Sozialforschung. In: Strobl R./Böttger A. (Hrsg.): Wahre Geschichten? Zur Theorie und Praxis qualitativer Interviews. Baden-Baden, 9-21
- Hopf C./Müller W. (1995): Zur Entwicklung der empirischen Sozialforschung in der Bundesrepublik Deutschland. In: Schäfers B. (Hrsg.): Soziologie in Deutschland. Opladen, 51-74
- Horstkemper M./Tillmann K.-J. (2003): Studien zu Einzelschulen. Erziehungswissenschaft, 14. Jg., H. 27, 45-78
- Huber A. (2001): Die Angst des Wissenschaftlers vor der Ästhetik. Forum Qualitative Sozialforschung (on-line) 2, No. 2
- Huxley T.H. (1870): Presidential Address at the British Association for "Biogenesis and Abiogenesis". Collected Essays, Vol. 8
- Jornitz S. (2008): Was bedeutet eigentlich „evidenzbasierte Bildungsforschung“? Die Deutsche Schule, 100 (2), 206-216
- Kalthoff H./Hirschauer S./Lindemann G. (Hrsg.) (2008): Theoretische Empirie: Zur Relevanz qualitativer Forschung. Frankfurt a. M.
- Kardorff E.v. (1995): Qualitative Sozialforschung – Versuch einer Standortbestimmung. In: Flick U./Kardorff E.v./Keupp H./Rosenstiel L.v./Wolff S. (Hrsg.): Handbuch Qualitative Sozialforschung. München, 3-9
- Keiner E. (2008): Fragen und Befunde der Wissenschaftsforschung der Erziehungswissenschaft. Vortrag auf dem 21. DGfE-Kongress „Kulturen der Bildung“ an der Universität Dresden, 17. März
- Kelle U. (2007). Die Integration qualitativer und quantitativer Methoden in der empirischen Sozialforschung. Wiesbaden
- Kelle U./Kluge S. (2006): Vom Einzelfall zum Typus. Fallvergleich und Fallkontrastierung in der qualitativen Sozialforschung. Wiesbaden
- Kerlinger F.N. (2000): Foundations of Behavioral Research. Orlando
- Kirchhöfer D. (1997): Qualitative versus quantitative Forschung – Schöpferische Spannung oder aufgehobene Dichotomie? In: Schmidt F. (Hrsg.): Methodische Probleme der empirischen Erziehungswissenschaft. Baltmannsweiler, 203-214
- Klauer K.J. (2001): Forschungsmethoden der Pädagogischen Psychologie. In: Krapp A./Weidenmann B. (Hrsg.): Pädagogische Psychologie, Weinheim, 75-98
- Klieme E./Bos W. (2000): Mathematikleistung und mathematischer Unterricht in Deutschland und Japan. Triangulation qualitativer und quantitativer Analysen am Beispiel der TIMS-Studie. Zeitschrift für Erziehungswissenschaft 3, 359-380
- Knoblauch H. (2000): Zukunft und Perspektiven qualitativer Forschung. In: Flick/Kardorff/Steinke (Hrsg.), 623-632
- Knoblauch H. (2008): Sinn und Subjektivität in der qualitativen Forschung. In: Kalthoff H./Hirschauer S./Lindemann G. (Hrsg.): Theoretische Empirie: Zur Relevanz qualitativer Forschung. Frankfurt a.M., 210-233
- König E./Zedler P. (Hrsg.) (1995): Bilanz qualitativer Forschung. Bd. 1: Grundlagen qualitativer Forschung. Weinheim
- Krammling-Jöhrens D. (1997): Atmosphäre als Wirklichkeitsebene. Eine ethnographische Studie über die Glocksee-Schule. Unveröff. Diss.: Fachbereich 1. Universität Gesamthochschule Kassel
- Krüger H.-H. (2000): Stichwort: Qualitative Forschung in der Erziehungswissenschaft. Zeitschrift für Erziehungswissenschaft 3, 323-342
- Kunter M./Schümer G./Artelt C./Baumert J./Klieme E./Neubrand M./Prenzel M./Schiefele U./Schneider W./Stanat P./Tillmann K.-J./Weiß M. (2002): PISA

- 2000: Dokumentation der Erhebungsinstrumente. Berlin: Max-Planck-Institut für Bildungsforschung
- Kvale S. (1995): Validierung. In: Flick U./Kardorff E. v./Keupp H./Rosenstiel L. v./Wolff S. (Hrsg.): Handbuch Qualitative Sozialforschung. München, 427-431
- Lamnek, S. (1995): Qualitative Sozialforschung. Band 1. Methodologie. Weinheim
- Lamnek S. (2005): Qualitative Sozialforschung. Band 2. Methoden und Techniken. Weinheim
- Leutner D. (1999): Hypothesenprüfung versus interpretative Exploration: Die endlose Debatte zur Funktion quantitativer und qualitativer Analysen in der Lehr-Lern-Forschung. Unterrichtswissenschaft, H. 4, 323-332
- Leutner D. (2000): Blick zurück nach vorn: Trends der Lehr-Lernforschung. Unterrichtswissenschaft 28, H. 1, 30-37
- Lochner R. (1963): Deutsche Erziehungswissenschaft. Meisenheim
- Lodge D. (1978): Changing Places. A tale of two campuses. London
- Lüders C. (2000): Herausforderungen qualitativer Forschung. In: Flick/Kardorff/Steinke (Hrsg.), 632-642
- Lüders C. (2003): Gütekriterien. In: Bohnsack R./Marotzki W./Meuser M. (Hrsg.): Hauptbegriffe qualitativer Sozialforschung. Opladen, 83-85
- Ludwig P.H. (1999): Imagination. Sich selbst erfüllende Vorstellungen zur Förderung von Lernprozessen. Opladen
- Ludwig P.H. (2001): Pädagogik als Disziplin mit kultur- und naturwissenschaftlichem Charakter. Eine unabhängig vom empirisch-analytischen Paradigma begründbare Klassifizierung. Pädagogische Rundschau 55, H. 1, 81-97
- Ludwig P.H. (2003): „Korrektiv gegenüber wildgewordener Theorie“. Zur Bedeutung der Empirie für die Pädagogik. Antrittsvorlesung an der Universität Kassel am 6. 2. 2002. Kasseler Universitätsschriften-Server (KUSS/Opus): <http://opus.uni-kassel.de/opus/volltexte/2003/58>
- Ludwig P.H. (2004): Zur Trennschärfe von sogenannten quantitativen und qualitativen Forschungsverfahren - Ein vereinfachter Definitionsvorschlag. Empirische Pädagogik 18, 337-349
- Ludwig P. H. (2007): 'Anything goes – quality stays'. Standards of quality for qualitative empirical research in education. Vortrag auf dem Kongress „Contested Qualities of Educational Research“ der „European Educational Research Association“ (EERA) an der Universität Gent, Belgien, am 21. Sept.
- Ludwig P. H. (2009): "Anything goes - quality stays". Quality standards for qualitative educational research in the context of justification. EducatiON-line. Database of the British Education Index: University of Leeds, UK [<http://www.leeds.ac.uk/educol/documents/180695.pdf>]
- Ludwig P.H./Ludwig H. (2006): Die Vergleichsperspektive in der einzelschulischen Qualitätsforschung. In: Ludwig H./Beutel S./Kleinespel K. (Hrsg.): Entwickeln - Forschen – Beraten. Weinheim, 180-198
- Mayring P. (1996): Einführung in die qualitative Sozialforschung. München
- Mayring P. (1999): Qualitativ orientierte Forschungsmethoden in der Unterrichtswissenschaft: Ein Anwendungsbeispiel aus der Lernstrategieforschung. Unterrichtswissenschaft, 292-309
- Mayring P. (2000): Qualitative Inhaltsanalyse. Weinheim
- Mecheril P. (2003): Prekäre Verhältnisse. Über natio-ethno-kulturelle (Mehrfach-) Zugehörigkeit. Münster
- Menck P. (2005): Gesichtspunkte für die Formulierung von Anträgen zur Förderung von Forschungsprojekten. Erziehungswissenschaft 16, H. 31, 6-15
- Meyer M. (2009): Abduktion, Induktion – Konfusion. Bemerkungen zur Logik der interpretativen Sozialforschung. Zeitschrift für Erziehungswissenschaft, 12, 302-320

- Moosburger H./Kelava A. (Hrsg.) (2007): Testtheorie und Fragebogenkonstruktion. Heidelberg
- Niegemann H. (2006): Lehr-Lern-Forschung. In: Rost D.H. (Hrsg.): Handwörterbuch Pädagogische Psychologie. Weinheim, 387-393
- Oswald H. (1997): Was heißt qualitativ forschen? In: Friebertshäuser B./Prengel A. (Hrsg.): Handbuch Qualitative Forschungsmethoden in der Erziehungswissenschaft. Weinheim, 71-87
- Popper K.R. (1966): Logik der Forschung. Tübingen
- Popper K. R. (1969): Die Logik der Sozialwissenschaften. In: Adorno T.W./Albert H./Dahrendorf R./Habermas J./Popper K. R. (Hrsg.): Der Positivismusstreit in der deutschen Soziologie. Neuwied, 103-124
- Popper K.R. (1992): Die offene Gesellschaft und ihre Feinde. Band I und II. Stuttgart
- Porter C. (1934): Anything Goes. Bühnenmusical (1936 verfilmt mit Bing Crosby)
- Prein G./Erzberger C. (2000): Integration statt Konfrontation! Ein Beitrag zur methodologischen Diskussion um den Stellenwert quantitativen und qualitativen Forschungshandelns. Zeitschrift für Erziehungswissenschaft 3, 343-357
- Preuss-Lausitz U. (1997): Geschlechtersozialisation und Schulpädagogik in der Nachmoderne. Die Deutsche Schule 89, 429-445
- Reichenbach H. (1938): Experience and prediction. An analysis of the foundations and the structure of knowledge. Chicago
- Reichenbach H. (1953): Der Aufstieg der wissenschaftlichen Philosophie. Berlin (orig.: The Rise of Scientific Philosophy. Berkeley 1951)
- Reicher S. (2000): Against methodolatry: Some comments on Elliott, Fischer, and Rennie. British Journal of Clinical Psychology, 39, 1-6
- Reichert J. (2000): Zur Gültigkeit von qualitativer Sozialforschung. Forum Qualitative Sozialforschung (on-line) 1, No. 2
- Reichert J. (2007): Qualitative Sozialforschung. Erwägen, Wissen, Ethik 18, 1-14
- Reinmann G. (2006): Nur „Forschung danach“? Vom faktischen und potentiellen Beitrag der Forschung zu alltagstauglichen Innovationen beim E-Learning. Arbeitsberichte Medienpädagogik. Augsburg
- Renkl A. (1999): Jenseits von $p < .05$: Ein Plädoyer für Qualitatives. Unterrichtswissenschaft, H. 4, 310-322
- Roeder B. (2006): Schülerbefragung. In: Rost D.H. (Hrsg.): Handwörterbuch Pädagogische Psychologie. Weinheim, 637-642
- Roeder P.M. (2003): TIMSS und PISA – Chancen eines neuen Anfangs in Bildungspolitik, -planung, -verwaltung und Unterricht. Zeitschrift für Pädagogik 49, 180-197
- Rombach H. (1965): Konfessionalität und Erziehungswissenschaft. Freiburg
- Rosenthal R. (1976): Experimenter effects in behavioral research. New York
- Rost D.H. (2005): Interpretation und Bewertung pädagogisch-psychologischer Studien. Weinheim
- Rost F. (2009): Letztlich zählt nur die Qualität. Statement zur Problematik wissenschaftlicher Online-Publikationen. Erziehungswissenschaft, 20. Jg., H. 38, 75-77
- Saldern M.v. (1995): Zum Verhältnis von qualitativen und quantitativen Methoden. In: König E./Zedler P. (Hrsg.): Bilanz qualitativer Forschung. Bd 1. Weinheim, 331-371
- Schnabel K. (1999): Wissenschaft ist Abstraktion, aber Abstraktes ist zu nix zu gebrauchen. Unterrichtswissenschaft, 333-336
- Schnaitmann G.W. (2004): Forschungsmethoden in der Erziehungswissenschaft. Zum Verhältnis von qualitativen und quantitativen Methoden in der Lernforschung an einem Beispiel der Lernstrategienforschung. Frankfurt a.M.
- Schwartz-Shea P. (2009): Judging Quality: Evaluative Criteria for Interpretive Empirical Research. Paper presented at the annual meeting of the American Political Science Association, Chicago, Sep 02, 2004 [access: Allacademic

- Research 2009-05-26 http://www.allacademic.com/meta/p61207_index.html;
Zugriff 20.2.2010]
- Seale C. (1999): The quality of qualitative research. London
- Slavin R.E. (2002): Evidence-based education policies: Transforming educational practice and research. *Educational Researcher* 31, No. 7, 15-21
- Slavin R.E. (2007): Evidence-based reform in Education. what will it take. Keynote address to the „European Conference on Educational Research“ at Ghent University (Belgium)
- Slavin R.E. (2008). Evidence-based reform in education: what will it take? *European Educational Research Journal* 7, No. 1, 124-128
- Smith J. K./Hodkinson P. (2005): Relativism, criteria, and politics. In: Denzin N./Lincoln Y. S. (eds): *The sage handbook of qualitative research*3. Thousand Oaks, CA, 915-932
- Soeffner H.-G. (1989): Anmerkungen zu gemeinsamen Standards standardisierter und nicht-standardisierter Verfahren in der Sozialforschung. In: Soeffner H.-G.: *Auslegung des Alltags – Der Alltag der Auslegung*. Frankfurt a.M., 51-65
- Steinke I. (1999): Kriterien qualitativer Forschung. *Ansätze zur Bewertung qualitativ-empirischer Sozialforschung*. Weinheim
- Steinke I. (2000): Gütekriterien qualitativer Forschung. In: Flick/Kardorff/Steinke (Hrsg.), 319-331
- Stelzl I. (2005): Fehler und Fallen der Statistik. Münster
- Terhart E. (1995): Kontrolle von Interpretationen: Validierungsprobleme. In: König/Zedler (Hrsg.), 373-297
- Uhl S. (1996): Die Mittel der Moralerziehung und ihre Wirksamkeit. Bad Heilbrunn
- Ulich D./Haußer K./Mayring P./Strehmel P. (1981): Kognitive Kontrolle in Krisensituationen: Arbeitslosigkeit bei Lehrern. Antrag an die Deutsche Forschungsgemeinschaft. Universität München
- Waider C. (2005): Evaluation von Outdoor-Trainings. Eine qualitative Untersuchung zu Wirkung und Transfer: In: Lakemann U. (Hrsg.): *Wirkungsimpulse von Erlebnispädagogik und Outdoor-Training*. Empirische Ergebnisse aus Fallstudien. Augsburg
- Wellenreuther M. (2000): *Quantitative Forschungsmethoden in der Erziehungswissenschaft*. Weinheim
- Wellenreuther M. (2005): *Lehren und Lernen – aber wie? Empirisch-experimentelle Forschungen zum Lehren und Lernen im Unterricht*. Baltmannsweiler
- Wernet A. (2006): *Hermeneutik – Kasuistik – Fallverstehen*. Stuttgart
- Wottawa H. (2006): Evaluation. In: Rost D. H. (Hrsg.): *Handwörterbuch Pädagogische Psychologie*. Weinheim, 162-168