

## Bildungspsychologische Forschung

Albert Ziegler

Dieser Beitrag basiert auf der unten im Kasten stehenden Definition von bildungspsychologischer Forschung. Offensichtlich lässt sich darin „Bildung“ durch Begriffe wie Denken, Gedächtnis oder Werbung ersetzen, ohne dass der Charakter einer wissenschaftlichen Definition verloren geht. Ihre besondere Pointe bezieht diese Definition somit aus dem Begriff „Bildung“. Gemeinsam mit dem Substantiv Psychologie bildet es die Nominalkomposition Bildungspsychologie, die eine Teildisziplin der Psychologie bezeichnet (Spiel & Reimann, 2005). Ihre Forschungsaktivitäten richten sich insbesondere auf die Entwicklung, Optimierung, Nutzung und Evaluation von Bildungsangeboten und auf individuelle Bildungsprozesse, ihre Charakteristika, Förderung, Ergebnisse und ihren Stellenwert für das Erreichen persönlicher Ziele.

Bildungspsychologische Forschung untersucht Phänomene der Bildung aus psychologischer Sicht. Es handelt sich dabei um eine systematische, methodengeleitete Suche nach neuen Erkenntnissen, deren Erfolg anhand von Qualitätsmaßstäben beurteilt werden kann.

### 1. Vorüberlegungen

Wissenschaftliche Teildisziplinen lassen sich in immanente und exoterische unterteilen. So sind beispielsweise Denk- und Gedächtnispsychologie immanente Teildisziplinen der Psychologie, da Denken und Gedächtnis theoretische Terme dieser Wissenschaft darstellen. Markt- und Werbepsychologie sind dagegen exoterische Teildisziplinen, da weder der Markt noch die Werbung theoretische Begriffe der Psychologie sind.

Immanente Teildisziplinen sind auf Fragestellungen innerhalb ihrer Mutterwissenschaft spezialisiert. Exoterische Teildisziplinen untersuchen aus der Perspektive der Mutterwissenschaft Fragestellungen, die außerhalb dieser angesiedelt sind. Ein Beispiel hierfür ist die Bildungspsychologie.

Auf den ersten Blick mag es verwundern, wieso sich überhaupt exoterische Teildisziplinen bilden können. Plausibel erschien es wohl eher, dass eine Wissenschaft Forschungsfragen ausschließlich aus sich selbst heraus hervorbringt und sich bei der Erklärung ihrer Forschungsbefunde vollständig auf ihre eigenen theoretischen Terme beschränkt. Im Folgenden werden daher sowohl extradisziplinäre als auch intradisziplinäre Gründe für die Etablierung einer exoterischen Teildisziplin vorgestellt.

#### 1.1 Extradisziplinäre Gründe für das Entstehen einer exoterischen Teildisziplin

Gesellschaften sind für ihr Funktionieren genauso auf Wissen angewiesen wie dies Individuen für das Erreichen ihrer Ziele sind. Das Bedürfnis nach informierten Entscheidungen steigt immer dann rasant an, wenn die Tragweite und Folgeschwere von Entscheidungen hoch ist und/oder falls es sich um immer wiederkehrende Probleme handelt. Ist in einer solchen Situation der Wissensbestand einer Wissenschaft geeignet, zur Entscheidungsoptimierung beizutragen, und lassen zudem künftige Forschungsaktivitäten dieser Wissenschaft weiteres relevantes Wissen erwarten, *kann* sich eine exoterische Teildisziplin herausbilden. So entdeckten beispielsweise in der Vergangenheit Werbetreibende, wie hilfreich psychologische Forschungsergebnisse sein können, und fragten in der Folgezeit gezielt weitere Forschungsergebnisse nach. Dies führte zur Etablierung der Werbepsychologie und ihres Forschungsapparats (Felsler, 2001).

Bekanntermaßen ist die Psychologie ein gefragter Auskunftgeber nicht nur in der Werbebranche, sondern in sehr unterschiedlichen Bereichen. Davon zeugt die Existenz spezieller Psychologien wie zum Beispiel Arbeits- und Pädagogische Psychologie, Gesundheits- und Rechtspsychologie, Medien- und Marktpsychologie, Religions- und Politische Psychologie, Umwelt- und Verkehrspsychologie. Die Aufzählung dieser exoterischen Teildisziplinen der Psychologie lässt einerseits vermuten, dass gesellschaftliche Funktionsbereiche und Handlungskontexte ein hohes Maß an Relevanz haben müssen, damit sie einer Wissenschaft genügend Anreize zur Spezialisierung bieten können. Sie veranschaulicht andererseits den enorm großen Anwendungsbereich psychologischen Wissens. Und in der Tat erscheint die Psychologie geeignet, gerade auch zum Thema Bildung wertvolle Beiträge zu leisten. Sie befasst sich wissenschaftlich fundiert mit vielfältigen Aspekten dieser Thematik. Insbesondere beruhen Bildungsprozesse auf Lernen, einem der Kernkonzepte der Allgemeinen Psychologie (Byrne, 2002). Die potentiellen Beiträge der Psychologie umfassen aber fast ihr gesamtes Kompetenzspektrum. So ist etwa Bildung durch Übung, Training oder Gestaltung anregender, informativer sozialer und räumlicher Kontexte gestaltbar. Personen unterscheiden sich in der Wahrnehmung, Aneignung und Nutzung von Bildungsangeboten. Es lassen sich entwicklungsabhängige Muster erkennen usw. Die Vielzahl der psychologischen Teildisziplinen, die unmittelbar wichtige Beiträge zum Verständnis und zur optimalen Steuerung von Bildungsprozessen leisten könnten, ist so groß, dass man durchaus von einer Querschnittsaufgabe für die gesamte Psychologie sprechen kann. Kurzum: Psychologie erscheint aus der Perspektive des Bildungssektors betrachtet schon jetzt als ein attraktiver Kooperationspartner, gezielte *bildungspsychologische* Forschungen versprechen sogar eine enorm gesteigerte Attraktivität.

## **1.2 Intradisziplinäre Gründe für das Entstehen einer exoterischen Teildisziplin**

Umgekehrt kann die Frage gestellt werden, weshalb eine Wissenschaft Interesse haben sollte, sich für extradisziplinäre Fragestellungen zu öffnen. Auf den vorliegenden Kontext gemünzt bedeutet dies, warum die Psychologie das Thema Bildung für derart bedeutungsvoll erachten sollte, dass sie dafür eigens eine neue Teildisziplin etabliert. Bei der Beantwortung dieser Frage spielen unter dem Gesichtspunkt von Forschung drei Gründe eine besondere Rolle: Forschungsinteressen, Forschungsmöglichkeiten und Gewinnung von Forschungsressourcen.

Eine neue Teildisziplin muss erstens weiten Teilen der Mutterdisziplin Forschungsfelder für vielfältige Fragestellungen bieten. Spricht sie ein Fach nicht in seiner Breite an, würde sie wahrscheinlich allenfalls als wichtige Spezialfragestellung einer bereits bestehenden Teildisziplin fungieren. Man kann mit Fug und Recht annehmen, dass ein Forschungsgegenstand *psychologische Aspekte der Bildung* sicherlich sehr viele Teildisziplinen der Psychologie anspricht, sodass er nicht mehr nur exklusiv von einer psychologischen Teildisziplin für sich reklamiert würde. So finden die Schwerpunktfächer der Psychologie wie Allgemeine Psychologie, Persönlichkeits-, Entwicklungs- und Sozialpsychologie, Psychologische Diagnostik beim Thema Bildung sicherlich genauso

fruchtbare und vielversprechende Forschungsfelder wie viele weitere Teildisziplinen, allen voran die Pädagogische Psychologie und mit etwas Abstand unter anderem Medienpsychologie, Klinische Psychologie und Organisationspsychologie. In der Tat belegt ein Blick in die wissenschaftlichen Zeitschriften, die sich schwerpunktmäßig mit bildungspsychologischen Themen befassen, wie bunt gefächert die intradisziplinäre Verortung der publizierenden Forscherinnen und Forscher in der Psychologie ist.

Eine potentielle Teildisziplin sollte gute Forschungsbedingungen und einen guten empirischen Zugang bieten. Auch dies ist bei der Thematik Bildung ohne Zweifel gegeben, da nicht nur ein passives, sondern sogar ein aktives Interesse an bildungspsychologischer Forschung besteht: (Bildungs-)Psychologinnen und -psychologen wird der Zugang nicht nur gewährt. Um ihr Engagement beispielsweise bei der Entwicklung von Unterrichtsprogrammen, Interventionen oder Evaluationen wird sogar geworben. Ferner sind Untersuchungen relativ leicht durchzuführen und Untersuchungsteilnehmerinnen und -teilnehmer recht einfach zu rekrutieren. Dem nicht zu leugnenden Handicap, dass in der Regel strikte experimentelle Kontrollen im Vergleich zu Laborstudien schwieriger zu realisieren sind, steht der Vorteil größerer ökologischer Validität gegenüber.

Neben Forschungsinteressen und Forschungsmöglichkeiten bildet die Möglichkeit, Forschungsressourcen für das eigene Fach zu gewinnen, einen weiteren gewichtigen intradisziplinären Grund für die Etablierung einer neuen Teildisziplin. Der Begriff Ressource ist hierbei weit gefasst und reicht von Forschungsgeldern bis hin zu Forschungsreputation (und der damit selbstverständlich verbundenen Möglichkeit, weitere Forschungsmittel zu akquirieren). Vor dem Hintergrund der gegenwärtig enormen Bedeutung des Themas Bildung, seiner Präsenz in Öffentlichkeit und Medien sowie den zur Zeit fließenden Forschungsmitteln besteht wohl Konsens, dass die Psychologie ein hohes Interesse daran haben sollte, sich dem Thema Bildung zuzuwenden.

## **2. Wissenschaftstheoretische Aspekte bildungspsychologischer Forschung**

Die Güte von Forschung ist Gegenstand der Wissenschaftstheorie. Ihre Aufgabe besteht darin, Forschungsaktivitäten einer Wissenschaft zu beschreiben und zu bewerten sowie im optimalen Fall, präskriptive Vorgaben zu machen. Wirft man jedoch als Bildungspsychologe einen Blick in die wissenschaftstheoretische Literatur, verwirrt diese mehr durch die Heterogenität der Standpunkte und Empfehlungen, als dass man Orientierung gewinnen könnte. So finden sich unter anderem die Ansätze Rationalismus (Aune, 1970; Cottingham, 1984) und Kritischer Rationalismus (Popper, 1968), Empirismus (Hume, 1748/1955; Kenny, 1986), logischer (Stove, 1973) und konstruktiver Empirismus (van Fraassen, 1980), Konstruktivismus (Ernest, 1994; Maturana & Varela, 1980) und Reismus (Lejewski, 1976; Woleński, 1996), Skeptizismus (Russel, 1948), Relativismus (Berger & Luckmann, 1966; Feyerabend, 1962), Funktionalismus (Fodor, 2003, Levin, 1985) und Pragmatismus (Aune, 1970; Wiener, 1990), Evolutionäre (Campbell, 1974) und Naturalistische Epistemologie (Quine, 1969) – um nur einige der Hauptansätze zu nennen.

Leider ist es nicht einmal in Ansätzen möglich, ein gemeinsames Extrakt dieser Wissenschaftstheorien zu benennen. Ohne einen metatheoretischen Rahmen kann eine Bildungspsychologie jedoch keine rationale Forschung leisten, weil dann beispielsweise Methodenfragen nicht entscheidbar wären. Es scheint, als müsse sie sich ihren methodologischen Rahmen selbst zimmern; in der Wissenschaftstheorie liegt er jedenfalls nicht bereit.

Ziel bildungspsychologischer Forschung ist die Erstellung eines Aussagesystems in Form einer wissenschaftlichen Theorie, das methodische Gütekriterien erfüllen muss (Gadenne, 1986). Die vordringliche Frage gilt dabei dem Wirklichkeitsausschnitt, auf den sich diese Forschung bezieht. Allerdings sieht man sich hier mit dem Problem konfrontiert,

dass der Forschungsgegenstand bildungspsychologischer Forschung nicht durch natürliche Begriffe wie Sonne, Baum oder Sauerstoff konstituiert wird. Er wird also nicht einfach vorgefunden, sondern Forscherinnen und Forscher müssen ihn konstruieren und erschließen.

Geht man plausiblerweise davon aus, dass der Gegenstand einer Wissenschaft identisch ist mit ihrem Forschungsgegenstand, dann finden sich bereits erste sehr brauchbare und diskutierbare Vorschläge in der metatheoretischen bildungspsychologischen Arbeit bei Spiel und Reimann (2005). Die Autorin und der Autor haben in dieser richtungweisenden Arbeit den Gegenstand der Bildungspsychologie anhand einer auf axiomatischen Vorüberlegungen beruhenden dreidimensionalen Listendefinition konstituiert. Ihr Strukturmodell bestimmt auf der zeitlichen Dimension Abschnitte der Bildungskarriere, die vorerst noch eng an westlichen Leitvorstellungen orientiert ist. Es werden weiter drei an systemtheoretische Vorläufer anknüpfende Handlungsebenen benannt sowie fünf Aufgabenbereiche. Neben Beratung, Prävention, Intervention sowie Bildungsmonitoring und Evaluation ist einer davon die Forschung selbst. Eine solche erste Ortsbestimmung des Forschungsgegenstands ist sicherlich noch fragil und bislang nicht praxiserprobt, doch sicherlich ein ganz ausgezeichneter erster Startpunkt.

### **3. Qualitätsmerkmale bildungspsychologischer Forschung**

Bildungspsychologische Forschung kann auf vielfältige Weise betrieben werden: Als Labor- und Feldforschung, experimentell und explorativ, im Längsschnitt oder im Querschnitt, als Einzelfallstudie oder in Form von Handlungsforschung, als qualitative oder quantitative Forschung (vgl. Lewis-Beck, Bryman & Liao, 2004). Im Prinzip bleibt es der Forscherin bzw. dem Forscher überlassen, welche Forschungsmethode am besten geeignet zur Erreichung der Forschungsziele erscheint. Doch gleichgültig, welche Methode gewählt wird, muss die Forschung bestimmten Qualitätsmaßstäben genügen (siehe auch Kerlinger, 1986). Die fünf zentralen Forderungen werden im Folgenden besprochen.

#### **3.1 Randomisierung**

Randomisierung erfolgt am besten in einem zweistufigen Prozess. Erstens sollten Personen zufällig ausgewählt und zweitens anschließend zufällig einer Bildungsbedingung zugewiesen werden. Ein solches Verfahren stellt bereits die interne Validität einer Studie sicher, das heißt die Ergebnisse sind zumindest für die Personengruppe der Stichprobe gültig (vgl. Shadish, Cook & Campbell, 2002). Allerdings muss die Randomisierung auch methodisch einwandfrei vorgenommen werden. Die Folgen eines Misslingens wurden erstmals durch die Ergebnisse einer der berühmtesten Studien der Wissenschaftsgeschichte, dem so genannten Lanarkshire Milchexperiment, der Forschergemeinde ins Bewusstsein gerufen (vgl. Student, 1931). Ziel der britischen Forschungsstudie war es festzustellen, ob die Verteilung von Milch positive Konsequenzen zeitigt. Obwohl die viele tausend Schülerinnen und Schüler umfassende Untersuchung sehr sorgfältig geplant war, darunter auch Zufallsauswahlen der teilnehmenden Schülerinnen und Schüler und Zufallszuweisungen zur Treatmentbedingung, entpuppten sich die Ergebnisse als wertlos. Lehrkräfte hatten die Randomisierung unterlaufen, indem sie die Milch nicht an die zufällig ausgewählten Schülerinnen und Schüler verteilten, sondern an die Bedürftigen.

Messergebnisse sind vielfältigen Einflüssen ausgesetzt. Um zu verhindern, dass diese systematisch in eine bestimmte Richtung beeinflussen, werden verschiedene Zufallszuweisungen, das heißt Randomisierungen vorgenommen.
--

Der zweistufige Randomisierungsprozess ist in der bildungspsychologischen Forschungspraxis oft nicht realisierbar. So können beispielsweise Schülerinnen und Schüler

nicht per Zufall bestimmten Schulformen wie Grund-/Primar-/Volksschule, der allgemeinbildenden höheren Schule/Gymnasium/Kantonsschule oder Realschulen zugewiesen werden. Dem stehen Hindernisse wie die Erfüllung von Leistungsvoraussetzungen, Alter, Bildungsentscheidungen der Erziehungsberechtigten etc. im Wege. In solchen Fällen muss man sich mit quasi-experimentellen Designs behelfen, das heißt auf die zufällige Zuweisung von Versuchspersonen zu Parallelgruppen verzichten.

Da quasi-experimentelle Designs die interne Validität gefährden und vermutlich beeinträchtigen, müssen bestimmte Vorsichtsmaßnahmen getroffen werden. So sollte in Prätests die Vergleichbarkeit der Treatmentbedingungen überprüft werden. Doch auch dies wird oft nicht realisierbar sein. Es muss daher von der Möglichkeit Gebrauch gemacht werden, unterschiedliche Voraussetzungen in der statistischen Auswertung zu kontrollieren, beispielsweise mit Hilfe so genannter Kovarianzanalysen. Allerdings ist es schlechterdings unmöglich, alle potentiell relevanten Ausgangsunterschiede zu berücksichtigen. Ferner sind die Möglichkeiten solcher Nachadjustierungen begrenzt (z.B. Pedhazur, 1997). Das Problem der Randomisierung kann daher als ein hartnäckiges und wahrscheinlich auch bleibendes Ärgernis bildungspsychologischer Forschung angesehen werden. Neben den erwähnten statistischen Behelfen wird man sich daher auf absehbare Zeit mit theoretischen Abschätzungen der verzerrenden Wirkung suboptimaler Randomisierung behelfen müssen.

### **3.2 Kontrolle**

Bildungspsychologisch relevante Prozesse finden in konkreten Handlungskontexten statt.

Spiel und Reimann (2005) unterscheiden dabei Mikro-, Meso- und Makroebene.

Beispielsweise kann die Mitarbeit einer Schülerin im Mathematikunterricht nicht allein durch in der unmittelbaren Situation wirkende Prozesse erklärt werden. Angemessene Erklärungen umfassen weitere Handlungsebenen wie Klasse, Schule, Familie, Gesellschaft und sogar Kultur. Den Einfluss all dieser Variablen in Experimenten zu kontrollieren, ist unmöglich.

Bildungspsychologische Forschung muss daher durch verschiedene Maßnahmen versuchen, den Einfluss von Störvariablen, die die Interpretation der Ergebnisse erschweren, möglichst gering zu halten. Die drei wichtigsten Strategien neben der statistischen Kontrolle (siehe 3.1) sind die Herstellung standardisierter Untersuchungsbedingungen, Selektion und Stratifizierung.

In Bezug auf empirische Studien meint der Begriff Kontrolle das Bemühen, den Einfluss der Randbedingungen einer Untersuchung zu minimieren.

Eine Standardisierung der Untersuchungsbedingungen – auch als Durchführungsobjektivität bezeichnet – hat offensichtlich die Folge, das „Rauschen“ in den Daten zu verringern. Als Rauschen erscheinen Bildungspsychologinnen und -psychologen dabei jegliche Fluktuationen in den Messungen, die sie sich nicht erklären können.

Es gibt aber auch systematische Einflüsse, welche die Daten bekanntermaßen beeinflussen, die jedoch aus irgendwelchen Gründen nicht erfasst werden können. Wenn man beispielsweise weiß, dass Elterneinstellungen Bildungsentscheidungen von Mädchen im mathematisch-naturwissenschaftlichen Bereich beeinflussen (siehe Schober, 2002), so könnten ausschließlich Mädchen von Eltern mit ähnlicher Einstellung in die Untersuchung aufgenommen werden. Hierdurch könnte der Einfluss elterlicher Einstellungen auf die Streuung der Daten minimiert werden. Wendet man hingegen alternativ die Strategie der Stratifizierung an, würden die Eltern anhand ihrer Einstellungen in verschiedene Gruppen eingeteilt. Der Einfluss ihrer Einstellungen würde somit bereits explizit bei der Gestaltung des Untersuchungsdesigns berücksichtigt. Bei der Auswertung der Ergebnisse können anschließend Aussagen zu den jeweiligen Elterngruppen gemacht werden. Der Vorteil dieser

Strategien besteht darin, dass sie Streuungen der Daten innerhalb der Untersuchungsgruppen verringern und somit die Wahrscheinlichkeit, einen Effekt zu finden, vergrößern.

### 3.3 Objektivität

Im letzten Abschnitt wurde bereits die Durchführungsobjektivität als strikte Standardisierung der Untersuchungsbedingungen erwähnt. Zusätzlich werden Auswertungs- und Interpretationsobjektivität unterschieden. Ihr Name deutet schon an, dass es keine Rolle spielen sollte, wer die Daten bildungspsychologischer Untersuchungen auswertet oder interpretiert.

Objektivität ist ein Güte Merkmal empirischer Forschung. Sie bezeichnet die Unabhängigkeit der Messergebnisse von den die Untersuchung durchführenden Forscherinnen und Forschern, Forschungsgruppen oder Forschungsinstitutionen.

Bei der Auswertung bildungspsychologischer Datensätze sind einige Grundsätze zu beachten, die leider immer wieder verletzt werden. So ist im Einklang mit dem von Spiel und Reimann (2005) vorgelegten Strukturmodell der Bildungspsychologie zu konstatieren, dass Bildung auf mehreren Handlungsebenen geschieht. Aus diesem Grund sollten für viele Fragestellungen die Auswertungseinheiten nicht nur auf einer Ebene angesiedelt sein (beispielsweise die Motivation von Schülerinnen und Schülern), sondern auch weitere Ebenen berücksichtigt werden (Motivation der Lehrkräfte, Schulklima, Eigenheiten des nationalen Bildungssystems etc.). Konkret könnte dies beispielsweise bedeuten, dass anstatt einfacher varianzanalytischer Designs hierarchische lineare Modelle (Bryck & Raudenbush, 1992) verwendet werden müssten. Ein weiteres Auswertungsproblem ist oftmals die unterschiedliche Bewertung von Rechenvoraussetzungen (z.B. Normalverteilungsannahme; Einfluss von Ausreißern, inhomogenen Varianzen und schiefen Verteilungen von Residuen) und die daraus resultierenden abweichenden Verwendungen von Auswertungsverfahren.

Wie gering die Interpretationsobjektivität ausfallen kann, zeigte sich jüngst bei der Bewertung der Ergebnisse der PISA-Studie. Obwohl im Prinzip allen Interpretinnen und Interpreten die gleichen Daten vorlagen, kamen sie zu höchst unterschiedlichen Bewertungen. Ein Ausweg könnte die Standardisierung von Interpretationen darstellen. Beispielsweise wird in vielen wissenschaftlichen Zeitschriften, die bildungspsychologische Arbeiten publizieren, neben einer Angabe der Irrtumswahrscheinlichkeit auch die Angabe von Effektstärken verlangt. Dies ist begrüßenswert. Die Standardisierung der Bewertungen, wie beispielsweise 0.5 sei eine mittlere Effektstärke (Cohen, 1988), ist dagegen häufig irreführend.

Aus der Not nur schwer herzustellender Interpretationsobjektivität lässt sich für die Bildungspsychologie zumindest eine Konsequenz ziehen und eine Hoffnung ableiten. Die Konsequenz besteht in der Forderung nach rigider Transparenz, sodass jederzeit nachvollziehbar ist, wie jemand zur eigenen Interpretation gelangte. Die Hoffnung besteht darin, dass abweichende Interpretationen keineswegs ein Unglück sein müssen, sondern auch Grundvoraussetzung einer fruchtbaren Diskussionskultur sind.

### 3.4 Reliabilität

Die Begriffsbestimmung von Reliabilität (s. u.) muss in zweierlei Hinsicht präzisiert werden. Gemeint ist nicht wie genau gemessen wird, also beispielsweise wie fein die Messskala unterteilt ist. Ob nun beispielsweise Schülerinnen bzw. Schüler dem oberen, mittleren und unteren Leistungsdrittel oder der oberen und unteren Leistungshälfte zugeordnet werden sollen, sagt daher noch nichts über die Messgüte aus. Wohl aber, ob diese Zuordnungen korrekt erfolgten. Weiter impliziert eine hohe Reliabilität keineswegs, dass tatsächlich die intendierte Größe gemessen wurde. Beispielsweise mag ein Leistungstest zwar sehr

zuverlässig messen, doch könnte es der Fall sein, dass er nicht nur Leistung, sondern auch Prüfungsängstlichkeit, Motivation und Konzentration erfasst.

Reliabilität ist ein Güte Merkmal empirischer Forschung. Sie bezeichnet die Zuverlässigkeit einer Messung, die umso höher ist, je geringer der Messfehler ist.

Es gibt verschiedene Techniken, die Reliabilität einer Messung abzuschätzen (vgl. Fisseni, 1990; Heller, 1984). Bildungspsychologische Studien sind dabei vor allem mit zwei Problemen konfrontiert. Erstens sind standardisierte Messinstrumente rar, weshalb beispielsweise zur Erfassung der Schulleistung die Reliabilität der verwendeten Messinstrumente oft nur an der Untersuchungsstichprobe abgeschätzt werden kann. Ein zweites Problem stellen die vielen hoch inferenten Messungen dar, also solche, in die Urteile und Schlussfolgerungen einfließen (Ingenkamp, 1997). Wenn beispielsweise bei der Unterrichtsbeobachtung oft nicht einmal genau abschätzbar ist, ob sich eine Schülerin gemeldet hat, dann kann man sich die Schwierigkeiten der Bildungsexpertinnen und -experten der OECD vorstellen, die nach einem wenige Tage umfassenden Besuch die nationale Bildungsqualität im Vorschulbereich abschätzen müssen. Die Möglichkeit ist nicht unplausibel, dass eine andere Delegation zu einem anderen Ergebnis gelangt wäre, da die Beantwortung einer solchen komplexen Fragestellung kaum auf der Basis von ausschließlich niedrig inferenten Messungen möglich ist.

### **3.5 Validität**

Validität ist ein Güte Merkmal empirischer Forschung. Sie bezeichnet die Gültigkeit einer Messung, das heißt ob tatsächlich die intendierte Größe erfasst wurde.

Messungen psychologischer Größen lassen sich nie direkt vornehmen. Sie bedürfen der Operationalisierung, das heißt der Überführung der Konstrukte in ein geeignetes Messverfahren. Die Validitätsproblematik lässt sich daher in zwei Teilfragestellungen aufspalten: 1) Was ist die beste Operationalisierung der zu messenden bildungspsychologischen Größe? 2) Wie gut ist die Operationalisierung gelungen?

Die erste Fragestellung nach der besten Operationalisierung ist zwar noch ungelöst, doch finden sich eine Menge brauchbarer Hinweise in einschlägigen Methodenhandbüchern (z.B. Babbie, 2007). Allerdings garantieren diese längst nicht Konsens. Ein prominentes aktuelles Beispiel eines Dissenses zwischen Bildungswissenschaftlerinnen bzw. -wissenschaftlern ist die unterschiedliche Auffassung bezüglich dessen, was in der internationalen Bildungsstudie PISA gemessen wird (Baumert, Brunner, Lüdtke & Trautwein, 2007; Rindermann, 2007). Obwohl niemand ernsthaft daran zweifelt, dass die PISA-Konstrukteurinnen und -Konstrukteure große Sorgfalt bei der Operationalisierung des Konstrukts Bildungsqualität walten ließen, wird kritisiert, dass gar nicht Bildungsqualität, sondern eigentlich Intelligenz gemessen wurde. Die zweite Fragestellung bezieht sich auf eine Abschätzung der erzielten Validität. Auch hier sind keine endgültigen Lösungen in Sicht, doch sehr hilfreiche Vorarbeiten verfügbar. Neben theoriegeleiteten Beurteilungen gibt es sogar quantitative Schätzungen in Form von Validitätskoeffizienten (Katzer, Cook & Crouch, 1998).

### **4. Fehlerquellen bildungspsychologischer Forschung**

Bildung und die mit ihr verbundenen Ziele und Methoden sind ein hoch brisantes politisches und weltanschauliches Thema. Bildungspsychologie ist daher wie kaum eine andere Wissenschaftsdisziplin der Gefahr von Verzerrungen und Irrtümern ausgesetzt. Aus diesem

Grund soll an dieser Stelle eine Systematik der fünf wichtigsten Fehlerquellen gegeben werden, die in Abbildung 1 grafisch zusammengefasst sind.

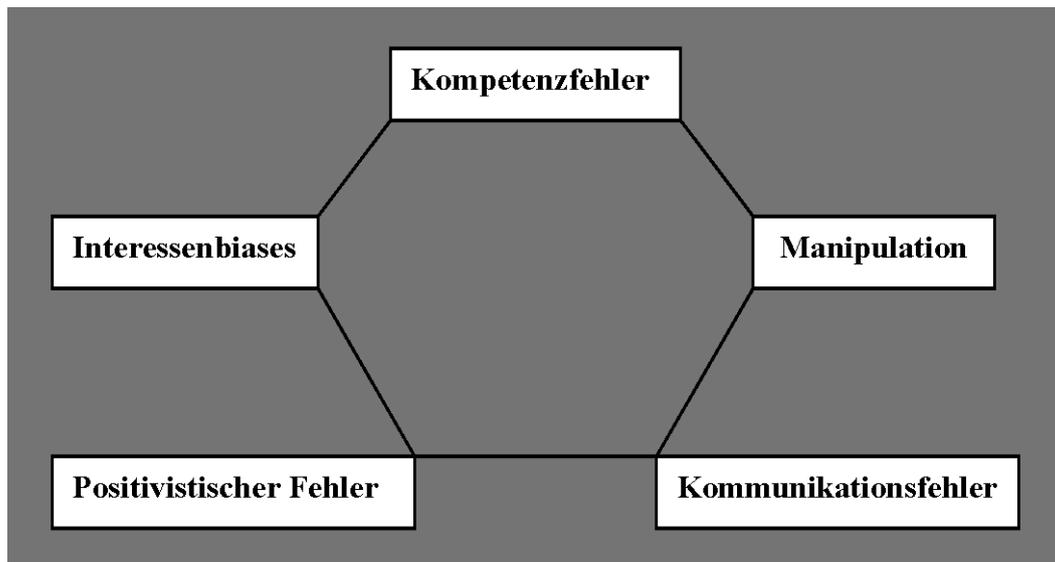


Abb. 1: Fehlerquellenpentagramm bildungspsychologischer Forschung

#### 4.1 Kompetenzfehler

Kompetenzfehler können auftreten, wenn Forscherinnen und Forscher allgemein akzeptierte Standards ihres Faches nicht beherrschen. Unkenntnisse können sich auf Theorien, Erhebungsmethoden, (statistische) Auswertungsverfahren etc. beziehen. Sie sollten keinesfalls als ein Ärgernis angesehen werden, das allenfalls zweitklassige Wissenschaftlerinnen und Wissenschaftler betrifft. Bereits eine Teildisziplin wie die Bildungspsychologie ist derart komplex, umfassend und informationsreich, dass keine Forscherin und kein Forscher mehr alle ihre Facetten gleichermaßen kompetent beherrscht. So kann beispielsweise die gängige Praxis des Peer-Reviews wissenschaftlicher Arbeiten mit der anschließend eingeräumten Möglichkeit ihrer Verbesserung als eine wissenschaftspolitische Praxis gewertet werden, Kompetenzfehler möglichst auszuschließen. Die im optimalen Fall recht elaborierten Fachgutachten geben der einzelnen Forscherin und dem einzelnen Forscher hilfreiche Hinweise zur Verbesserung seiner Arbeit.

Viele weithin verbreitete Kompetenzmängel haben das Interesse der Forschungsgemeinde stimuliert und führten zu einer bunten Ratgeberkultur. Zwei der bekanntesten Fallen, in die Forscherinnen und Forscher immer wieder hineintappen, sollen im Folgenden kurz vorgestellt werden. Der *ökologische Fehler* beruht darauf, dass Gruppendaten überinterpretiert werden. Seinen Namen verdankt er William Robinson, der in einer Bildungsstudie auf ihn stieß (Robinson, 1950). Untersuchungsgegenstand waren die Lesefertigkeiten in den damals 48 Bundesstaaten der USA. Dabei wurden Immigrantinnen und Immigranten mit in Amerika geborenen Bürgerinnen und Bürgern verglichen. Der Hauptbefund war ein deutlicher Zusammenhang zwischen der Lesefertigkeit und der Anzahl von Immigrantinnen und Immigranten in einem Bundesstaat. Robinson warnte jedoch davor, die Korrelation von 0.53 so zu interpretieren, dass Immigrantinnen und Immigranten über eine höhere Lesefertigkeit verfügten. Tatsächlich lag die Korrelation auf individueller Ebene bei -0.11! Es stellte sich heraus, dass Immigrantinnen und Immigranten verstärkt dort siedelten, wo die Lesefertigkeit durchschnittlich höher war. Robinson warnte davor, Schlussfolgerungen bezüglich Individuen auf der Basis von Gruppendaten zu ziehen.

Ein weiterer verbreiteter Fehler ist die kausale Interpretation korrelativer Zusammenhänge (Hooke, 1983). Wenn zwei Variablen P und Q korrelieren, so kann dies

zumindest vier Gründe haben: 1) P verursacht Q, 2) Q verursacht P, 3) beide sind durch eine weitere Variable verursacht oder 4) die Korrelation war zufällig. So trivial dies scheint, werden in der Praxis doch viele Fehler begangen. Den hohen statistischen Zusammenhang von verkaufter Eiscreme und Anzahl Ertrunkener an Stränden werden die Wenigsten als Beleg für die Gefährlichkeit von Eiscreme interpretieren. Ganz offensichtlich ist die Anzahl von Personen an den Stränden Grund dieser Korrelation. Allerdings kann gerade die Häufigkeit als Drittvariable die Interpretation vieler bildungspsychologischer Befunde erschweren. Beispielsweise wurde das Bild von Hochbegabten als „Problemkinder“ stark durch Untersuchungen geprägt, die an Hochbegabtenberatungsstellen durchgeführt wurden (Freeman, 1979; Grassinger, 2007). Es ist jedoch offensichtlich, dass in dieser Stichprobe Hochbegabte mit Problemen überrepräsentiert sind.

#### **4.2 Manipulation**

Manipulation meint den bewussten Versuch, die Ergebnisse und die Rezeption von Forschungsstudien zu steuern. Theoretisch kann dies auf sehr vielfältige Weise geschehen. Beispielsweise können in Interviews die Antworten durch geschickte Formulierung der Fragen in eine bestimmte Richtung gelenkt werden. Nachdem die Daten vorliegen, kann ihre Rezeption durch eine selektive Auswahl, unangemessene statistische Verfahren oder verzerrte Darstellungen, beispielsweise in Form irreführender Grafiken, manipuliert werden (Tufté, 1997).

Es ist naiv zu hoffen, die sich gerade etablierende Bildungspsychologie würde frei sein von Manipulationsversuchen. So besteht kein Grund anzunehmen, dass sich in ihr aus irgendeinem Grund weniger schwarze Schafe tummeln sollten als es gemeinhin in der Wissenschaft der Fall ist (Jaffe & Spierer, 1987). Im Gegenteil steht zu erwarten, dass die Politisierung und Ideologisierung von Bildung zu vielerlei Beeinflussungsversuchen führt, gegenüber denen sich nicht alle Bildungspsychologinnen und -psychologen als resistent erweisen werden. So wird sicherlich manche Institution die Vergabe von weiteren Forschungsgeldern an die Willfährigkeit bildungspsychologischer Forscherinnen und Forscher knüpfen.

Gegenmaßnahmen sind erstens eine Wissenschaftsgemeinde, die Transparenz und Nachvollziehbarkeit als Forschungstugend einfordert. So sind beispielsweise die Messinstrumente und die Daten der internationalen PISA-Studie allen Forscherinnen und Forschern zugänglich. Zweitens sollten zentrale bildungspsychologische Forschungsarbeiten obligatorisch ein Peerreviewsystem durchlaufen, bei dem sie von kompetenten Fachvertreterinnen und -vertretern begutachtet werden. Drittens sollten Replikationsstudien durchgeführt werden. Diese drei Gegenmaßnahmen sind bereits heute zumindest in Ansätzen in der Bildungspsychologie realisiert. Die vierte Gegenmaßnahme beruht auf der Bildung einer eigenen Fachvertretung mit den ihr eigenen Möglichkeiten, insbesondere jenen zur Bildung von Ethikkommissionen und zur Aussprechung von Sanktionen.

#### **4.3 Kommunikationsfehler**

Die Ergebnisse bildungspsychologischer Forschung werden Fachkolleginnen und -kollegen sowie der Öffentlichkeit kommuniziert. Idealerweise werden sämtliche Ergebnisse mitgeteilt. In der Praxis ist es jedoch häufig so, dass nur eine Auswahl weitergegeben wird. Dies ist beispielsweise dann gerechtfertigt, wenn Störungen während der Datenerhebung auftreten oder sich Daten als Messfehler behaftet herausstellen. Inakzeptabel ist jedoch beispielsweise der so genannte Publikationsbias (auch als „file-drawer problem“ bezeichnet). Er bezeichnet die Tendenz, vor allem statistisch signifikante Befunde zu berichten, also nur solche Studien, bei denen „etwas herauskam“. Bei einem Signifikanzniveau von 5% tritt unter 200 Studien jedoch ohnehin bei ca. 10 ein *statistisch* signifikanter Effekt auf, ohne dass dies einen realen Effekt widerspiegelt.

Die Entscheidung für oder gegen eine Publikation von Daten hat gravierende Folgen. So zeigte etwa die Metaanalyse von Lipsey und Wilson (1993), dass pädagogische Interventionen eine bereinigte durchschnittliche Effektstärke von .47 Standardabweichungen aufweisen. Bei publizierten Daten ist die durchschnittliche Effektstärke jedoch .14 Standardabweichungen größer, so dass fast ein Viertel der berichteten Wirkungen pädagogischer Interventionen auf eine verzerrte Kommunikation empirischer Daten zurückzuführen ist.

Ein Ausweg für die Bildungspsychologie könnte eine Entscheidung ähnlich des im September 2004 verkündeten Beschlusses mehrerer renommierter medizinischer Fachzeitschriften sein. Sie publizieren von der Pharmaindustrie gesponserte Studien nur unter der Bedingung, dass sie vor Beginn in einer öffentlichen Datenbank registriert wurden. Dies soll garantieren, dass alle Daten publiziert werden. Ebenso wäre es beispielsweise wünschenswert, wenn insbesondere bei bildungspsychologischen Evaluationsstudien sämtliche erfassten Variablen vor Beginn der Untersuchungen angezeigt würden.

#### **4.4 Positivistischer Fehler**

Bildungspsychologische Forschung findet nicht im luftleeren Raum statt, sondern wird stets innerhalb bestimmter Gegebenheiten durchgeführt. Das Vorhandene<sup>1</sup> bildet dabei sowohl den bevorzugten Forschungsgegenstand als auch den Forschungsrahmen und den bevorzugten Referenzwert.

In Deutschland werden Studien zur Sekundarstufe meist mit Schülerinnen und Schülern durchgeführt, die eine Schulform in einem mehrgliedrigen Schulsystem besuchen. Dagegen werden in den Vereinigten Staaten während dieses Bildungsabschnitts nur Schülerinnen und Schüler eines ungegliederten Schulsystems untersucht. Man kann also salopp formuliert nur das untersuchen, was vorhanden ist.

Das Positive bildet auch den Forschungsrahmen. Erweist sich beispielsweise eine bestimmte pädagogische Maßnahme an Montessorischulen als ineffektiv, kann daraus nicht geschlossen werden, dass es nicht doch eine Schulform gäbe, an der genau diese pädagogische Maßnahme effektiv wäre.

Bei der Beurteilung der Effektivität wird aber fast immer – wie auch im zuletzt angeführten Beispiel – das Positive als Referenzwert gewählt. Dieser hat unter Umständen eine große Auswirkung auf die Beurteilung der Effektivität einer bildungspsychologischen Maßnahme. Beispielsweise könnte sie im Vergleich zum finnischen Schulsystem als ineffektiv, im Vergleich zum deutschen beziehungsweise österreichischen Schulsystem als effektiv bewertet werden, weil die durchschnittlichen Schulleistungen in diesen Ländern voneinander abweichen.

#### **4.5 Interessenbiases**

Forscherinnen und Forscher wählen ihre wissenschaftlichen Fragestellungen selektiv aus. Dies hat vielfältige Ursachen, die teilweise durchaus plausibel sind. Beispielsweise werden Bildungspraktikerinnen und -praktiker in so unterschiedlichen Ländern wie Korea, Saudi-Arabien und der Schweiz sehr verschiedene Wissensbedürfnisse haben, die sie an die Bildungspsychologie herantragen. Allerdings kann es auch der Fall sein, dass Bildungspsychologinnen und -psychologen signalisiert bekommen, manche Ergebnisse seien erwünschter als andere. Eindrucksvolle Beispiele finden sich in Gould (1996), der akribisch dokumentierte, wie sich in einem xenophoben und ethnophoben gesellschaftlichen Klima

---

<sup>1</sup> Im Anschluss an die wissenschaftstheoretische Position des Positivismus wird das Vorhandene (wertneutral) als das Positive bezeichnet.

Bildungspsychologinnen und -psychologen nur schwer Erwartungen widersetzen konnten und erwünschte Forschungsergebnisse produzierten.

Ferner stehen Forscherinnen und Forscher ihren eigenen Ergebnissen nicht neutral und unparteiisch gegenüber. Sie favorisieren bestimmte über andere, sei es, weil diese ihre Hypothesen bestätigen, ihrer Weltanschauung oder politischen Meinung besser entsprechen, weitere Forschungsmittel an bestimmte Resultate geknüpft sind etc. Unterschiedliche Interessen beeinflussen ebenfalls die verwendeten Methoden und die Ergebnisse (z.B. Habermas, 1968; Gould, 1996) auf verschiedene Weise (Evans, 1989; Jaffe & Spirer, 1987).

Grundkenntnisse über Interessenwirkungen in den Erhebungssituationen sind mittlerweile fester und unverzichtbarer Bestandteil des methodischen Ausbildungskanons künftiger Sozialwissenschaftlerinnen und -wissenschaftler. Ihr genaues Verständnis ist umso wichtiger, da sie zumeist unterhalb der Bewusstseinschwelle von Forscherinnen und Forschern sowie Untersuchungsteilnehmerinnen und -teilnehmern wirken. Beispielsweise existiert bei Untersuchungsteilnehmerinnen und -teilnehmern die Tendenz zu erraten, was mit der Studie bezweckt wird und welche Ergebnisse erwünscht sind. Ein nicht geringer Anteil verhält sich dann konsistent mit diesen Vermutungen. Ein anderer Interessenbias beruht auf dem Wunsch von Personen, ein möglichst positives Bild der eigenen Person zu vermitteln. Insbesondere in Evaluationsstudien kann dies für Bildungspsychologinnen und -psychologen ein schwierig zu überwindendes Hindernis auf dem Weg zu validen Daten darstellen. Diese beiden Beispiele für Interessenbiases in der Erhebungssituation betrafen Untersuchungsteilnehmerinnen und -teilnehmer. Natürlich ist jedoch auch die Forscherin bzw. der Forscher nicht vor Interessenbiases gefeit. Beispielsweise kann in einer Befragungsstudie die Interviewerin bzw. der Interviewer durch Tonfall und Mimik eine Präferenz für eine bestimmte zur Auswahl stehende Antwortalternative verraten. Bei manchen Antworten wird sie bzw. er vielleicht zufriedener schauen als bei anderen. Wenn die Untersuchungsteilnehmerinnen und -teilnehmer im weiteren Interviewverlauf in diese erwünschte Richtung antworten, kann dies auch ein reiner Versuchsleiterinnen- bzw. Versuchsleitereffekt sein.

## **5. Bildungspsychologische Forschung als Chance**

Im letzten Abschnitt wurden mögliche Fehler- und Gefahrenquellen für bildungspsychologische Forschung beleuchtet. Aber natürlich bietet die Etablierung neuer Teildisziplinen stets auch Chancen für die Forschungen der Mutterwissenschaft und benachbarter Wissenschaftsdisziplinen. Insbesondere durch neue Fragestellungen und Perspektiven, durch innovative methodische Zugänge und Analyseverfahren kann es zu fruchtbaren Rückkoppelungen kommen. Bildungspsychologie erscheint als Impulsgeberin aus mehreren Gründen prädestiniert.

Verschiedene Forscherinnen und Forscher (z. B. Shotter, 2005; St. Julien, 2005) halten gegenwärtige Konzeptualisierungen von Lern- und Bildungsprozessen für inadäquat und kritisieren insbesondere, dass deren dynamische Komplexität nicht ausreichend abgebildet und genügend verstanden wird. Tatsächlich werden Lern- und Bildungsprozesse meist auf der Basis eines veralteten Forschungsverständnisses untersucht. Psychologie, Soziologie, Kultur- und Sozialwissenschaften sind auch heute noch leicht identifizierbare Kinder des Zeitalters der Aufklärung: Sie sind den Ideen der exakten Berechenbarkeit, Prognostizierbarkeit, Reduzierbarkeit und linear-mechanischer Prozesse verpflichtet (Mainzer, 2004; Starobinski, 2003). In der Tat war nach einem Bonmot von Kauffman (1995) die Wissenschaft des 18. Jahrhunderts als Folge der Newtonschen Revolution eine Wissenschaft organisierter Einfachheit. Die Wissenschaft des 19. Jahrhunderts befasste sich mit Hilfe der statistischen Mechanik mit desorganisierter Komplexität. Erst im 20. Jahrhundert befasste sie sich mit organisierter Komplexität. Doch die wissenschaftliche Orientierung an der alten Physik als

Theorie und ihrer Technologie als Praxis ist für die Bildungspsychologie ungeeignet (Midgley, 2004).

Bildung ist offensichtlich polytelisch, das heißt es gibt nicht eine einzige beste Bildung, sondern unzählige gleichberechtigte individuelle Möglichkeiten. Jörg, Davis und Nickmans (2007) schreiben: "We are already forced to acknowledge that we simply cannot anticipate the type of competences that will define for example a "literate" or "numerate" citizen a few decades from now. Indeed, we cannot even know if contemporary notions of literacy and numeracy will make sense 20 years into the future." Es ist in den Sozialwissenschaften nicht klar, was der konzeptuelle Analyserahmen, beispielsweise mathematisch und statistisch, für eine solche Offenheit in der Zielsetzung sein könnte. Eine explizite Akzeptanz von Polytelie sprengt beispielsweise den traditionellen empirischen Erziehungsbegriff der Pädagogik (vgl. Brezinka, 1977), deren Anliegen – vereinfacht ausgedrückt – die Entwicklung geeigneter Erziehungstechniken zur Erreichung fixer Erziehungsziele war. Doch in welche Richtung kann die konzeptuelle Weiterentwicklung gehen, die von bildungspsychologischer Forschung angestoßen werden könnte?

In Bildungsprozessen, die sich über ein gesamtes individuelles Leben erstrecken, bauen Individuen aus einfachsten Anfängen ein komplexes Handlungsrepertoire auf. Komplexität bedeutet in diesem Fall, dass Bildungsprozesse durch Vernetzungen unterschiedlichster Art charakterisiert sind. Im Strukturmodell der Bildungspsychologie nach Spiel und Reimann (2005) kommt dies beispielsweise durch das Postulat hierarchischer Handlungsebenen zum Ausdruck. Beispielsweise sind die Akteurinnen und Akteure Mitglieder von Familie, Bildungsinstitutionen, vielfältigen sozialen Gruppen, ihrer Gesellschaft etc. Diese Entitäten beziehungsweise Systeme und ihre Teile verändern sich und ihre Beziehungen untereinander permanent (Kelso, 1995). Ein besser geeigneter mathematischer Ansatz zur Erforschung von Bildungsprozessen wäre deshalb beispielsweise die mathematische Katastrophentheorie (Reno, 1989), die es auch gestattet, systemische Wechselwirkungen zu berücksichtigen.

Ein weiteres Kennzeichen von Bildungsprozessen ist, dass die Handlungen ihrer Akteurinnen und Akteure selbstreferentiell sind, das heißt Handlungen wirken in vielfacher Weise auf sich selbst zurück. Beispielsweise kann der Erwerb einer Lernstrategie künftige Bildungsprozesse beschleunigen. Die Erfassung von Selbstreferentialität impliziert daher zwingend die Berücksichtigung von Veränderungsprozessen in Längsschnittdesigns. Ein erfreulicher Einfluss bildungspsychologischer Forschung könnte somit darin bestehen, dass sie die Notwendigkeit der Berücksichtigung der zeitlichen Perspektive aufweist und Schrittmacher für die Etablierung raffinierter Forschungsdesigns wird.

Die in der bildungspsychologischen Forschung betrachteten zeitlichen Veränderungen und insbesondere die selbstreferentiellen Prozesse verlaufen keineswegs linear (Kauffman, 1991). Diese Erkenntnis, die in Bezug auf kritische Lebensereignisse, Krisen, Übertritte oder auch individuelle Entscheidungen trivial erscheint, hat noch längst nicht zur Entwicklung nonlinearer Analyseverfahren geführt. Im Gegenteil, gegenwärtig sind die in den Sozial- und Humanwissenschaften dominierenden mathematischen und statistischen Analyseverfahren wie Regressions- und Varianzanalysen, Strukturgleichungsmodelle oder hierarchisch lineare Modellierung dem algebraischen Spezialfall linearer Prozesse verpflichtet. Es wäre ein enormer Verdienst bildungspsychologischer Forschung, den theoretischen und methodischen Blick durch Beachtung von Nonlinearität zu erweitern und durch die Berücksichtigung von Polytelie, selbst-organisierter Komplexität und nicht zuletzt systemtheoretischer Ansätze zu bereichern.

## **6 Weiterführende Literatur**

Babbie, E. (2007). *The practice of social research* (11. ed.). Belmont, CA: Wadsworth.

- Creswell, J.W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2. ed.). Thousand Oaks, CA: Sage.
- Heller, K.A. (1991). *Begabungsdagnostik in der Schul- und Erziehungsberatung*. Bern: Huber.
- Ingenkamp, K. & Lissman, U. (2005). *Lehrbuch der pädagogischen Diagnostik*. Weinheim: Beltz.
- Vogt, W.P. (2005). *Dictionary of statistics and methodology: A nontechnical guide for the social sciences* (3. ed.). Newbury Park, CA: Sage.

## 7. Literatur

- Aune, B.A. (1970). *Rationalism, empiricism, and pragmatism*. New York: Random House.
- Babbie, E. (2007). *The practice of social research* (11. ed.). Belmont, CA: Wadsworth.
- Baumert, J., Brunner, M., Lüdtke, O. & Trautwein, U. (2007). Was messen internationale Schulleistungsstudien? Resultate kumulativer Wissenserwerbsprozesse. Eine Antwort auf Heiner Rindermann. *Psychologische Rundschau*, 58, 118-128.
- Berger, P.L. & Luckmann, T. (1966). *The social construction of reality: A treatise in the sociology of knowledge*. Garden City, NY: Anchor Books.
- Brezinka, W. (1977). *Grundbegriffe der Erziehungswissenschaft: Analyse, Kritik, Vorschläge*. München: Ernst Reinhardt.
- Bryck, A.S. & Raudenbush, S.W. (1992). *Hierarchical linear models: applications and data analysis methods*. Newbury Park: Sage Publications.
- Byrne, J.F. (Ed.). (2002). *Learning and memory* (2. ed.). Farmington Hills, Mi.: Macmillan.
- Campbell, D. (1974). Evolutionary epistemology. In P.A. Schilpp (Ed.), *The philosophy of Karl Popper* (pp. 413-463). La Salle, IL: Open Court.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. ed.). New York, NY: Academic Press.
- Cottingham, J. (1984). *Rationalism*. London: Paladin Books.
- Ernest, P. (Ed.) (1994). *Constructing mathematical knowledge: Epistemology and mathematics education*. London: Falmer.
- Evans, J.St.B.T. (1989). *Bias in human reasoning: Causes and consequences*. Hillsdale, NJ: Erlbaum.
- Felser, G. (2001). *Werbe- und Konsumentenpsychologie* (2. Aufl.). Heidelberg: Spektrum.
- Feyerabend, P. (1962). Explanation, reduction and empiricism. In H. Feigl & G. Maxwell (Eds.), *Minnesota Studies in the Philosophy of Science*, Vol. III (pp. 28-97). Minneapolis, MI: University of Minnesota Press.
- Fisseni, H.-J. (1990). *Lehrbuch der psychologischen Diagnostik*. Hogrefe: Göttingen.
- Fodor, J.A. (2003). *Hume variation*. Oxford: Oxford University Press.
- Freeman, J. (1979). *Gifted children*. Lancaster, UK: MTP.
- Gadenne, V. (1986). Darstellung und Prüfbarkeit psychologischer Theorien aus der Sicht der Aussagenkonzeption und des Non-Statement-View. In M. Amelang (Hrsg.), *Bericht über den 35. Kongreß der Deutschen Gesellschaft für Psychologie in Heidelberg 1986* (S. 191-198). Göttingen: Hogrefe.
- Gould, S.J. (1996). *The mismeasure of man*. New York: Norton.
- Grassinger, R. (2007). *Entwicklung und Evaluation eines systemischen Beratungsmodells für hochbegabte Kinder und Jugendliche auf der Basis des Aktiotop-Ansatzes* (unveröffentlichte Dissertation). München: Ludwig-Maximilians-Universität.
- Habermas, J. (1968). *Erkenntnis und Interesse*. Frankfurt a.M.: Suhrkamp.
- Heller, K.A. (1984). *Leistungsdiagnostik in der Schule*. Bern: Huber.
- Hooke, R. (1983). *How to tell the liars from the statisticians*. New York: Dekker.

- Hume, D. (1748/1955). *An inquiry concerning human understanding*. Indianapolis, IN: Bobbs-Merrill.
- Ingenkamp, K. (1997). *Lehrbuch der pädagogischen Diagnostik* (4. Aufl.). Weinheim: Beltz.
- Jaffe, A.J. & Spierer, H.F. (1987). *Misused statistics*. New York: Dekker.
- Jörg, T., Davis, B. & Nickmans, G. (2007). Towards a new, complexity science of learning and education. *Educational Research Review*, 2 (2), 145-156.
- Katzer, J., Cook, K.H. & Crouch, W.W. (1998). *Evaluating Information: A Guide for Users of Social Science Research* (4. ed.). Boston, MA: McGraw-Hill.
- Kauffman, S. (1991). Antichaos and adaptation. *Scientific American*, 265, 78-84.
- Kauffman, S. (1995). *At home in the universe. The search for the laws of self-organization and complexity*. Oxford: Oxford University Press.
- Kelso, J.A. (1995). *Dynamic Patterns: The self-organization of brain and behaviour*. Cambridge, MA: MIT Press.
- Kenny, A. (1986). *Rationalism, Empiricism and Idealism*. Oxford: Oxford University Press.
- Kerlinger, F.N. (1986). *Foundations of behavioral research* (3. ed.). New York: Holt, Rinehart & Winston.
- Lejewski, C. (1976). Outline of ontology. *Bulletin of the John Rylands University Library of Manchester*, 59, 127-147.
- Levin, J. (1985). Functionalism and the argument from conceivability. *Canadian Journal of Philosophy*, 11, 85-104.
- Lewis-Beck, M.S., Bryman, A. & Liao, T.F. (Ed.). (2004). *The Sage encyclopedia of social science research methods*. Thousand Oaks, CA: Sage.
- Lipsey, M.W. & Wilson, D.B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist*, 48, 1181-1209.
- Mainzer, K. (2004). *Thinking in complexity. The computational dynamics of matter, mind, and mankind* (4. ed.). Berlin: Springer.
- Maturana, H. & Varela, F. (1980). *Autopoiesis and cognition: The realization of the living*. Boston, MA: Reidel.
- Midgley, M. (2004). *The myths we live by*. London: Routledge.
- Pedhazur, E.J. (1997). *Multiple regression in behavioral research* (3. ed.). Fort Worth: Harcourt Brace.
- Popper, K. (1968). *The logic of scientific discovery* (3. ed.). London: Hutchinson.
- Quine, W.V. (1969). *Ontological relativity and other essays*. New York: Columbia University Press.
- Reno, T. (1989). *Structural stability and morphogenesis: An outline of a general theory of models*. Reading, MA: Addison-Wesley.
- Rindermann, H. (2007). The g-factor of international cognitive ability comparisons: The homogeneity of results in PISA, TIMSS, PIRLS and IQ-tests across nations. *European Journal of Personality*, 21, 667-706.
- Robinson, W.S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review*, 15, 351-357.
- Russell, B. (1948). *Human knowledge: Its scope and limits*. New York: Simon and Schuster.
- Schober, B. (2002). *Entwicklung und Evaluation des Münchner Motivationstrainings (MMT)*. Regensburg: Roderer.
- Shadish, W.R., Cook, T.D. & Campbell, D.T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton, Mifflin and Company.
- Shotter, J. (2005). Bringing corporeal life back in: chiasmic relations and poetic understanding. In W.E. Doll, M.J. Fleener, D. Trueit & J. St. Julien (Eds.), *Chaos, complexity, curriculum, and culture: a conversation* (pp. 195-208). New York: Peter Lang.

- Spiel, C. & Reimann, R. (2005). Bildungspsychologie. *Psychologische Rundschau*, 56, 291-301.
- Starobinski, J. (2003). *Action and reaction. The life and adventures of a couple*. New York: Zone Books.
- St. Julien, J. (2005). Complexity: Developing a more useful analytic for education. In W.E. Doll, M.J. Fleener, D. Trueit & J. St. Julien (Eds.), *Chaos, complexity, curriculum, and culture: a conversation* (pp. 101–116). New York: Peter Lang.
- Stove, D. (1973). *Probability and Hume's inductive scepticism*. Oxford: Clarendon.
- Student (1931). The Lanarkshire Milk Experiment. *Biometrika*, 23, 398-406.
- Tufte, E.R. (1997). *Visual explanations: Images, evidence and narrative*. Cheshire, CT: Graphics Press.
- van Fraassen, B.C. (1980). *The scientific image*. Oxford: Oxford University Press 1980.
- Wiener, P.P. (Ed.) (1990). *Charles S. Peirce, selected writings: Values in a universe of chance*. New York: Dover.
- Woleński, J. (1996). Reism in the Brentanist tradition. In L. Albertazzi et al. (Eds.), *The school of Franz Brentano* (pp. 357-375). Dordrecht: Kluwer.

### **Schlagworte**

Forschung, Wissenschaftstheorie, Randomisierung, Kontrolle, Objektivität, Reliabilität, Validität, Kompetenzfehler, Manipulation, Kommunikationsfehler, Positivistischer Fehler, Interessenbiases