

**Moral und Bildung**  
**Eine Kritik von Kohlbergs Theorie der moralisch-**  
**kognitiven Entwicklung**

**Georg Lind**

Zweite, korrigierte Auflage  
1998

**Copyright © 1998 by Georg Lind.**

Jede Verwendung von Teilen dieses Textes (mit Ausnahme von kurzen Zitaten bis zu 30 Worten) in Papier- oder anderen Medien bedürfen der schriftlichen Genehmigung durch den Copyright-Inhaber.

Der Text darf für rein private Zwecke und für Zwecke im Rahmen *öffentlicher* Bildung (kostenlose Skripte) frei ausgedruckt und vervielfältigt werden. Eine kommerzielle Nutzung dieses Textes oder Teilen davon bedürfen grundsätzlich der Genehmigung durch den Copyright-Inhaber.

Die erste Auflage dieses Buches war 1993 im Roland Asanger-Verlag, Heidelberg, erschienen.

"Die einzig wahrhafte Kraft gegen das Prinzip von Auschwitz wäre  
Autonomie, wenn ich den Kantischen Ausdruck verwenden darf, die Kraft  
zur Reflexion, zur Selbstbestimmung, zum Nicht-Mitmachen."

Theodor W. Adorno

Für  
Gisela und Gregor

## Inhalt

0	Vorwort zur zweiten Auflage .....	i-iv
1	Zu einer Bildungstheorie der Moralentwicklung .....	9
2	Das Zwei-Aspekte-Modell des moralischen Urteilsverhaltens und der Moralentwicklung .....	31
3	Kohlbergs <i>Moral Judgment Interview</i> – Eine Kritik .....	65
4	Der <i>Moralisches Urteil-Test</i> - Ein experimenteller Fragebogen .....	79
5	Moralische Bildung versus moralische Reifung: Das Problem der Regression .....	95
6	Bildung versus Sozialisation I: Kann moralische Urteilsfähigkeit wie eine politische Einstellung simuliert werden? .....	115
7	Bildung versus Sozialisation II: Kulturelle Einflüsse auf die moralische Urteilsfähigkeit .....	127
8	Förderung der moralischen Urteilsfähigkeit durch Bildungsprozesse: Ein Interventionsexperiment .....	153
9	Bedingungen für den Effekt von Interventionen zur Förderung der moralischen Urteilsfähigkeit: Eine Meta-Analyse .....	187
10	Moralische Stereotype und pluralistische Ignoranz: Die Wahr- nehmung moralischer Bedürfnisse von Schülern durch Lehrer .....	209
11	Bildungserfahrungen als Bedingung für moralische Entwicklung ...	239
12	Epilog .....	257
	Literaturverzeichnis .....	271
	Anhang: <i>Moralische Atmosphäre Fragebogen (MAF)</i> .....	291

## Danksagung

Ich möchte den folgenden Personen und Institutionen für unterschiedliche Formen der Unterstützung bei der Durchführung und Darstellung dieser Untersuchungen danken: Tino Bargel, der *Deutsche Forschungsgemeinschaft* (Bonn), Peter Dobbstein-Osthoff, *EMNID* (Bielefeld), der *Forschungsgruppe Gesellschaft und Region e.V.* (Konstanz), Dr. Monika Kuhn, dem *Kultusministerium von Nordrhein-Westfalen*, Ottfried Lind, Leonore Link, Manfred Scholz, der *Universität Konstanz*, Christa Kolbert-Ramm, John Sandberger, Dr. Heinz Schirp, Karin Tritt.

Die erste Auflage wurde gedruckt mit Unterstützung des Förderungs- und Beihilfefonds Wissenschaft der VG Wort.

## 0 Vorwort zur zweiten Auflage

Diese zweite Auflage von *Moral und Bildung* war lange überfällig. Die erste (nicht sehr große) Auflage im Jahr 1993 im Asanger-Verlag, Heidelberg, war innerhalb kurzer Zeit vergriffen. Aber kein Verlag sah eine Chance, mit einer zweiten Auflage den Erfolg der ersten zu wiederholen.

Dabei einige Rezensenten, wie der Göttinger Moralpsychologe und Pädagoge Günter Schreiner, das Buch als „einen Meilenstein im moralpädagogischen Diskurs“ gelobt. Roland Wakenhut, Professor für Sozial- und Organisationspsychologie an der Katholischen Universität Eichstätt, schrieb über das Buch gar: „Für die moralpsychologische Forschung, die nach dem Tode von Lawrence Kohlberg und der Auflösung des Harvard Center for Moral Education ihren Schwerpunkt nach Europa verlagert hat, liefert die Arbeit von Dr. Lind wichtige Impulse.“ Der Schweizer Erziehungswissenschaftler Fritz Oser (Universität Fribourg) schrieb in seinem Gutachten zu *Moral und Bildung* u.a.: "Selbst wenn man nicht die gleichen theoretischen Positionen wie Georg Lind vertritt, muß man zugeben, daß er einen der bedeutendsten Beiträge zur Moralforschung im deutschen Sprachraum leistet und geleistet hat.“

Über diesem Buch zugrundeliegenden, neuen Forschungsmethoden haben internationale Resonanz gefunden. Über die von mir vorgeschlagene Methodologie zur „Messung“ von kognitiven Strukturen im Bereich des moralischen Denkens und Handelns, dem „Experimental Questionnaire“ (EQ) schreibt Peter H. Rossi (University of Massachusetts, Amherst):<sup>a</sup> „I was delighted to learn of your EQ method. Your method has a lot more theory behind it than we have put behind the idea of the Factorial Survey and, with your permission, I would borrow some of your ideas. [. . .] One of the special feature of the Factorial Survey method is the idea of sampling factorial universes. [This idea] makes it possible to employ many dimensions and many levels within dimensions, providing the opportunity to study domains of considerable richness and texture. It also means abandoning ANOVA approaches and using multiple regression methods, with subsequent loss of ability to study higher order interactions among dimensions.“ Dr. John M. Bynner (The Open University, Milton Keynes, UK) schreibt ebenfalls 1984: „I think your arguments about attitude measurement are extremely persuasive. They tie with the general case for what Lamiell makes for ‚idiothetic measurement‘.“

*Moral und Bildung* ist ein wissenschaftlicher Text. Er faßt die Forschung des Autors über einen Zeitraum von fast zwanzig Jahre zusammen und schlägt auf der Grundlage dieser Forschung eine neue Theorie des moralischen Verhaltens und der Moralentwicklung vor. Viele Teile stellen hohe Anforderungen an das Vorverständnis der Leser. Gleichwohl habe ich versucht, den Text so verständlich wie möglich zu machen. Insbesondere die Einleitungen zu den einzelnen Kapitel sind in einer allgemein verständlichen Sprache gehalten.

Zudem handelt der Text von einem - je nach Standpunkt - altmodisch-konservativ erscheinenden Thema. Zwar beklagen viele einen angeblichen oder wirklichen „Verlust an moralischen Werten“ und rufen nach „mehr Moral“ im Privatleben, im nachbarlichen Zusammenleben und in der Politik. Aber, so glaubt man, Wissenschaft (zumal eine der Aufklärung verpflichtete) kann hierzu keinen Beitrag leisten. Selbst im (deutschsprachigen) Wissenschaftsbetrieb gilt das Thema Moral - von einigen Ausnahmen abgesehen - bei vielen als heikel, wenn nicht gar als anrüchig. Im Gegensatz zu vielen anderen Ländern, befassen sich hier nur wenige experimentell arbeitende Sozialwissenschaftler mit diesem Thema. So verwundert es auch kaum, daß das Buch (wie man an der geringen Zahl von Rezensionen und Zitationen sieht) von vielen deutschsprachigen Lehrbuchautoren (die einzige gründliche Rezension ist in englisch erschienen) bislang ignoriert wurde. Eine englischsprachige Auflage von *Moral und Bildung* muß ich leider auf einen unbestimmten Zeitpunkt verschieben. Auszüge und Erweiterungen des in *Moral und Bildung* vorgelegten *Zwei-Aspekte-Modells* der moralisch-kognitiven Entwicklung und der *Bildungstheorie* der Moral liegen jedoch als englische Artikel und Vorträge vor und sind im Internet zugänglich (<http://www.uni-konstanz.de/ag-moral/>).

Die Überarbeitung der zweiten Auflage von *Moral und Bildung* beschränkt sich auf die Korrektur von Tippfehlern, Auslassungen und stilistischen Holprigkeiten. Ich habe versucht, die Seitenzahlen und die Nummerierung der Tabellen und Graphiken soweit als möglich so zu belassen wie in der ersten Auflage, so daß das Zitieren und Querverweisen nicht allzusehr erschwert wird. Die Seitenzahl im Inhaltsverzeichnis beziehen sich auf die erste, gedruckte Auflage. Sie müssen nicht mit den Seitenzahlen hier übereinstimmen.

In manchen Fällen sind die Abbildungen auf eine andere Seite geraten, was dem „Fortschritt“ der elektronischen Textverarbeitung zu „verdanken“



ist. Bedauerlicherweise war es in vielen Fällen nicht möglich, die Graphiken elektronisch in das Format der Version 8 von Corel WordPerfect zu konvertieren, obwohl dasselbe Textsystem verwendet wurde wie bei der Erstellung der Druckvorlage der ersten Auflage (dort allerdings Version 5.1).

Sach- und Personenregister sowie Hyperlinks konnten in dieser Auflage noch nicht eingefügt werden. Sie sind für eine spätere Neuauflage geplant.

Abschließend möchte ich den Kollegen in der Universitätsbibliothek Konstanz, insbesondere Herrn Paul Kuhn, herzlich danken, daß sie diese zweite elektronische Auflage möglich gemacht haben.

Hinweise, Kritik und Nachfragen an die Adresse des Autors sind willkommen:

PD Dr. Georg Lind  
Universität Konstanz  
Sozialwissenschaftlichen Fakultät  
D-78457 Konstanz  
E-mail: [Georg.Lind@uni-konstanz.de](mailto:Georg.Lind@uni-konstanz.de)

Konstanz, April 1998

Der Autor

## 1 Zu einer Bildungstheorie der Moralentwicklung

Die Entwicklung der moralischen Urteilsfähigkeit wird in einem ganz entscheidenden Ausmaß von Bildungserfahrungen gefördert, das heißt, sie ist weder das Ergebnis eines sozialen Anpassungsdrucks, noch ist sie das Ergebnis eines "natürlichen" Reifungsprozesses, der einer inneren Entwicklungslogik folgt. Die für die Moralentwicklung notwendigen Bildungserfahrungen werden am ehesten von den allgemeinbildenden Schulen und Hochschulen bereitgestellt werden. Diese These, Kern einer *Bildungstheorie* der Moralentwicklung, steht im Zentrum der hier vorgelegten Untersuchung. Diese These ist nicht neu, aber sie wurde bislang kaum empirisch untersucht. Noch weniger wurde bislang mit Daten belegt, ob sie gegenüber den konkurrierenden und weitaus populärerem Theorien der Moralentwicklung bestehen kann, zum Beispiel neben der "Sozialisationstheorie", wonach die Entwicklung der Moral das Ergebnis eines sozialen Anpassungsdrucks ist, oder der "Reifungs- oder Emergenztheorie", nach der sich moralische Urteilsfähigkeit mit dem Alter quasi von allein, ohne oder sogar entgegen äußerer Einflüsse, als Ergebnis eines genetisch gesteuerten Reifungsprozesses entfaltet.

Nach der *Sozialisationstheorie* und der sozialen Lerntheorie geschieht Moralentwicklung durch das Erlernen von Normen und Werten der Gruppen, denen ein Mensch angehört. Einer der "Väter" der Sozialisationstheorie war der französische Soziologe und Erziehungstheoretiker Emile Durkheim. Für ihn lebt der Mensch von Natur aus in einem unstabilen Zustand, bedrängt von seinen Trieben, die es zu begrenzen und im Sinne der Gesellschaft umzulenken gilt. Dies bedarf, so Durkheim (1984/1902), des Einsichtigmachens sozialer Normen und Werte durch Zwang und Autorität. Ziel der Moralentwicklung ist nach dieser Theorie der Erwerb moralischer Einstellungen, an dem, in neueren Versionen, kognitive Prozesse beteiligt sein können. Nicht thematisiert - und von einigen Vertretern dieser Theorie sogar als überflüssig angesehen - werden die kognitiven Fähigkeiten, die notwendig sind, um moralische Werte oder Prinzipien in situationsgemäßes Verhalten zu übersetzen und die moralischen Konflikte zu lösen, die dabei auftreten können.

Nach der *Reifungstheorie* dagegen ist der Mensch von Geburt moralisch gut oder bekommt die Anlagen fürs Gut-sein mit; die Gesellschaft hingegen ist die Wurzel des Bösen und stellt mit ihren Normen und Zwängen eine Behinderung der Moralentwicklung dar. Das Spektrum der Reifungstheore-

tiker der Moral ist breit, dennoch gibt es einige gemeinsame Überzeugungen. "Alles ist gut, wie es aus den Händen des Schöpfers kommt; alles entartet unter den Händen des Menschen", leitet einer der großen Vertreter dieser Denkschule, J.-J. Rousseau (1971/1762, S. 9), sein epochales Werk *Emile oder Über die Erziehung* ein. Er insistiert darauf, "daß der Mensch von Natur aus gut ist..., [aber] die Gesellschaft den Menschen verdirbt und widernatürlich macht" (S. 241). Sigmund Freud und seine psychoanalytische Schule wiederholen die Rousseausche These in abgewandelter Form: Die herrschenden gesellschaftlichen Normen führten zur Unterdrückung menschlicher Triebe; diese Triebunterdrückung sei mitverantwortlich für unmoralisches Verhalten. Die Psychoanalyse vertraut daher darauf, daß, wie Arno Plack (1969) resümiert, "die menschliche Natur, wenn sie in Ruhe gelassen wird und nicht von klein auf verbogen, aus sich heraus ein Verhalten entwickelt, das dem Individuum wie der Gesellschaft am besten entspricht" (S. 345). Infolgedessen kann auch institutionalisierte Bildung keine positiven Effekte für die individuelle Moralentwicklung haben und kann es der Pädagogik, wie es von Braunmühl (1975) auf den Punkt brachte, nur um ihre eigene Abschaffung gehen.

Die kognitive Entwicklungstheorie der Moralentwicklung von J. Piaget und L. Kohlberg nimmt unter den bisherigen Ansätzen eine Sonderstellung ein. Sie weist Elemente beider Theorie auf. "Die Entwicklung kognitiver Strukturen sind", so Kohlberg (1984), "eher das Resultat von Vorgängen der Interaktion zwischen der Struktur des Organismus und der Struktur der Umwelt als das direkte Resultat der Reifung" (S. 8). Aber Richtung und Abfolge der Entwicklung sind, nach dieser Theorie, (genetisch?) festgelegt. Kohlberg nimmt unter Verweis auf Piaget an, daß die Stufen der Moralentwicklung „eine invariante Abfolge, Ordnung bzw. Aufeinanderfolge in der individuellen Entwicklung“ bilden (S. 14). Während kulturelle Faktoren diese Entwicklung beschleunigen, verlangsamen oder anhalten können, können sie nicht ihre Abfolge ändern. Auch die Prognose der kognitiven Entwicklungstheorie ist optimistisch. Statt 'der Mensch *ist* gut', nimmt diese Theorie an, 'der Mensch *wird* gut im Verlaufe seiner kognitiv-moralischen Entwicklung'. Zumindest wird er nicht schlechter. Aufgrund der inneren "Logik" der Moralentwicklung können soziale Einflüsse die Moralentwicklung nur beschleunigen oder verlangsamen, aber sie können sie nicht umkehren. Diese Annahme der *invarianten Sequenz* oder Nicht-Regression

bildet einen zentralen Pfeiler der Theorie und beherrscht auch, im Gegensatz zu manchen ihrer Aussagen über die Rolle der sozialen Umwelt, die Forschung von Kohlberg und seinen Kollegen. Ihre Theorie ist also im Kern eine Reifungstheorie.

Zahlreiche Studien der letzten Jahre haben einerseits Kohlbergs (1969) scharfe Kritik sozialisationstheoretischer Ansätze wie auch naiv-reifungstheoretischer Ansätze untermauert und scheinbar auch Anhaltspunkte für eine Bestätigung der kognitiven Entwicklungstheorie erbracht. Andererseits haben sie aber auch Hinweise dafür erbracht, daß einige dieser Bestätigungen auf unzulänglichen Forschungsdesigns und teilweise voreingenommenen Forschungsmethoden beruhen (Kapitel 3). Hauptaufgabe der hier vorgelegten Untersuchung war daher, diese Mängel - auf der Basis der kognitiven Entwicklungstheorie - zu beheben. Das Ergebnis zeigt, daß auch Kohlbergs Theorie die moralische Entwicklung nicht adäquat beschreibt. Vielmehr ist dafür eine Theorie notwendig, die wir oben als "Bildungstheorie" der Moralentwicklung bezeichnet haben.

### **1.1 Praktische Bedeutung**

Die vorliegende Untersuchung ist primär grundlagenorientiert; sie will eine Entscheidung zwischen drei alternativen Erklärungsansätzen herbeiführen. Das schließt nicht aus, daß diese Erklärungsansätze oder Thesen praktische Bedeutung haben, vor allem für pädagogische, bildungspolitische und sozialtherapeutische Maßnahmen. Im Gegenteil, die Frage nach der Rolle, die Bildungsprozesse für die Moralentwicklung haben, scheint von höchster Relevanz für die Ziele und Methoden jeder Moral- oder Sozialerziehung, für Erziehung überhaupt. Eine Bestätigung der Emergenzthese würde bedeuten, daß Kinder am besten zu moralisch reiferen Wesen werden, wenn man sie sich selbst überläßt und sie möglichst wenig erzieht. Tatsächlich haben viele "Reform-Pädagogen" und "Anti-Pädagogen" wie Alexander S. Neill (1969) in seiner berühmten Summerhill-Schule ihre Methoden durch Hinweis auf die "natürliche" Entwicklung moralischer Fähigkeiten begründet. Die Bildungspolitik könnte sich, sofern diese Sicht empirisch belegbar wäre, auf die Bereitstellung von Freiräumen - im konkreten und im übertragenen Sinne des Wortes - beschränken. Wenn Kohlbergs kognitive Entwicklungstheorie zutrifft, wäre das schlimmste, was bei fehlender Förderung passieren kann,

eine Verlangsamung oder Stagnation der Moralentwicklung. Was aber, wenn diese Theorie nicht zutrifft, wenn sich die Moralentwicklung bei fehlender Förderung umkehrt? Wir werden zeigen, daß dies entgegen der Kohlberg-Theorie, tatsächlich der Fall ist. Wenn die pädagogische Stimulation durch allgemeinbildende Institutionen (in der Regel die Schule) wegfällt, nimmt die moralische Urteilsfähigkeit der Menschen deutlich ab (Kapitel 5).

Eine Bestätigung der Sozialisationstheorie würde das genaue Gegenteil bedeuten. Es würde heißen, daß die Erziehungsinstanzen – Familie, Schule – ihr besonderes Augenmerk auf die Vermittlung von sozialen Normen und auf die Kontrolle ihrer Einhaltung legen müßten, um aus dem zunächst amoralischen Kind ein moralisch kompetentes Mitglied der Gesellschaft zu machen. Der Heranwachsende brauchte die moralischen Normen nur insoweit zu verstehen, wie ihm dies das Einhalten derselben erleichtert (Durkheim 1976/1924; Brezinka 1988). Dagegen sagt die Bildungstheorie der Moralentwicklung voraus, daß es mehr bedarf als das Verstehen sozialer Normen und der Anpassung hieran, daß für die Moralentwicklung vielmehr, wie Piaget und Kohlberg annahmen, entwickelte kognitive Fähigkeiten wesentlich sind. Des weiteren sagt diese Theorie voraus, daß die Forderungen, die sich aus der Sozialisationstheorie ableiten lassen, weitgehend ins Leere laufen, da fast alle Kinder, selbst in sehr jungen Jahren, eine hohe Sensibilität für grundlegende moralische Prinzipien haben und diese ihnen nicht erst vermittelt werden muß.<sup>1</sup> An was es Kindern - und auch vielen Jugendlichen und Erwachsenen - mangelt, ist die Fähigkeit, moralische Prinzipien in konkreten Situationen anzuwenden und moralische Konflikte zwischen widerstreitenden Prinzipien zu lösen.

Die Bildungstheorie impliziert daher einen dritten Weg der Erziehung. Sie besagt, daß Moral ein Aufgabenfeld von ähnlicher Struktur und mit ähnlichen Erfordernissen darstellt wie jedes andere Aufgabenfeld, dem wir uns im Leben stellen müssen, und daß daher ganz ähnliche didaktische Bemühungen notwendig sind, um moralische Urteilsfähigkeit optimal zu fördern. Moralische Probleme können natürlich nicht mit denselben Methoden und Kompetenzen gelöst werden wie mathematische oder physikalische Probleme. Aber auch hier gilt: Es müssen a) Kenntnisse ("moralisches

1 Neuere Untersuchungen zur Moralentwicklung im Kindheitsalter werden in Kapitel 2 referiert; s. auch Lind 1989a.

Wissen") vermittelt und b) Möglichkeiten für die Erprobung dieser Kenntnisse an praktischen Aufgaben geschaffen werden. Wirksame Moralerziehung erfordert eine spezielle Didaktik, gut ausgebildete Pädagogen und entsprechende Lehrpläne. Erste Ansätze hierfür liegen vor.<sup>2</sup> Freilich muß hier auch an Kohlbergs Warnung vor dem "psychologischen Fehlschluß" erinnert werden. Eine psychologische Theorie mag noch so wahr, empirisch getestet und vollkommen sein, sie kann keinen umfassenden Handlungsplan für alle Akteure in einem Anwendungsfeld anbieten. Handeln und Entscheiden in lebensweltlichen Bezügen ist mehr als die Anwendung einer Theorie, es ist die Koordination einer Vielfalt von Wissensformen und Theorien unterschiedlicher Herkunft in einem bestimmten Augenblick in einer bestimmten Situation von einer bestimmten Person. Es wäre falsch, die umfassende Handlungskompetenz, die Menschen, selbst Kinder, in solche Entscheidungen einbringen, durch eine einzige, monolithische Theorie über einen relativ kleinen Aspekt der Lebenswelt ersetzen zu wollen. Der Moralpsychologe ist immer auf die Zusammenarbeit mit den Betroffenen "im Feld", mit den Lehrern, Eltern und Kindern angewiesen.

Nach der Bildungstheorie der Moralentwicklung sind demokratisch-diskursive Prozesse nicht nur eine zusätzliche Bedingung für die Entwicklung der moralischen Urteilsfähigkeit, sondern eine unverzichtbare Voraussetzung (s. Kapitel 8 und 9). Die hier vorgelegten Ergebnisse untermauern die zentrale Bedeutung, die der moralischen Urteilsfähigkeit im individuellen Entwicklungsprozeß zukommt. Die Tatsache, daß ausgesprochen "kognitive" Interventionen, die Denken und Diskussionen über moralische Konflikte zum Inhalt haben, einen konsistenten Effekt von beachtlicher Größe erbringen, spricht eher für die Gültigkeit der kognitiven Entwicklungstheorie als für Reifungs- und Sozialisationstheorien zur Erklärung der Veränderungen im Moralverhalten von Jugendlichen und jungen Erwachsenen. Der Befund, daß Bildungsinstitutionen, aber nicht die berufliche Umwelt, aufs Ganze gesehen einen stark fördernden Einfluß auf die Moralentwicklung haben, spricht für die Bedeutung von föhigkeitsfördernden Maßnahmen für die Moralentwicklung, und gegen Reifungs- oder Anpassungsthesen und deren theoretische und

2 Power et al. 1989; Lind & Raschert 1987; Oser & Althof 1989.

praktische Implikationen.<sup>3</sup> Entwicklungsfördernde, moralische Diskurse können prinzipiell überall stattfinden, also auch in der Familie, im Freundeskreis, am Arbeitsplatz, finden aber offenbar die günstigsten Voraussetzungen in allgemeinbildenden Schulen. Wenn die Moralentwicklung nicht durch Bildungsprozesse ausreichend stimuliert wird, stagniert die Entwicklung der moralischen Urteilsfähigkeit oder bildet sich sogar zurück. Jugendliche, die das Schulsystem früh verlassen und in die Berufswelt eintreten, in der weniger Raum für dialogische Prozesse ist, weisen eine deutliche *Abnahme* der moralischen Urteilsfähigkeit auf (Kapitel 5). Kohlberg und viele von ihm inspirierte Pädagogen haben gezeigt, daß die Schulen ihre Aufgabe, die Förderung moralisch-demokratischer Kompetenzen, besser als bisher erfüllen könnten, wenn sie sich gezielt und methodisch darum bemühen würden (Kapitel 8 und 9). Daß die Förderung solcher Kompetenzen heute, angesichts einer komplexen Umwelt mit ihren vielfältigen Problemen dringlich geboten ist, wird von Philosophen wie Jonas (1979), Böhler (1992) und Mittelstraß (1992) mit überzeugenden Argumenten untermauert.

Warum ist dann die Bedeutung von Bildungserfahrungen in der Moralpsychologie bislang fast gänzlich übersehen worden? Neben der Tatsache, daß weder in der Sozialisation- noch in der Reifungstheorie Platz für die "Variable" Bildung ist, scheint mir der Grund darin zu liegen, daß Bildungserfahrungen heute so allgemein und für viele so selbstverständlich sind, daß uns ihre Effekte kaum mehr bewußt werden. In vielen Gesellschaften besteht eine langjährige Schulpflicht, so daß die Effekte von Bildungserfahrungen heute nur noch schwer von rein altersbedingten (Reifung-) Prozessen unterschieden werden können. Was in Wirklichkeit bildungsbedingte Entwicklungen sind, wird häufig einer altersabhängigen Reifung zugeschrieben.

3 Diese These steht nicht in Widerspruch zu den Untersuchungen von Lempert (1988) und seinen Mitarbeitern, die bei den von ihnen interviewten Industriearbeitern Fortschritte in der Moralentwicklung gefunden haben. Auch am Arbeitsplatz können Bildungsgelegenheiten vorhanden sein.

## 1.2 Definition und Operationalisierung

Nachdem deutlich geworden ist, daß eine Klärung der empirischen Gültigkeit der drei hier behandelten Theorien der Moralentwicklung nicht bloß von akademischem Interesse ist, sondern auch von höchst praktischem, stellt sich die Frage, wie sich eine empirische bzw. experimentelle Bewertung herbeiführen läßt. Eine Voraussetzung dafür ist offenkundig, daß diese Theorien einen tatsächlichen, von jedem beobachtbaren Sachverhalt betreffen, so daß sich anhand intersubjektiver Prüfung ein Konsens über Richtigkeit oder Falschheit erreichen läßt. Die Forderung nach intersubjektiver Überprüfbarkeit klingt einfacher als sie ist. James B. Watson (1970/1924), ein Begründer der modernen Psychologie, sah sie dadurch erfüllt, daß sich jede Theorie "auf jene Dinge beschränkt, die beobachtet werden können, und daß nur bezüglich dieser Dinge Gesetze formuliert werden." Beobachtet werden kann, so Watson, "Verhalten – was der Organismus *tut oder sagt*" (S. 6). Nur wenn Theorien sich an diesen - trivial erscheinenden - Grundsatz halten, können sie intersubjektiv nachvollzogen und anhand empirischer Daten beurteilt werden. Leider wurde dieser methodologische Grundsatz von seinen frühen Vertretern, den *Behavioristen* ad absurdum geführt, da sie - weit über das Ziel hinauschießend - *a priori* festlegten, daß all jene psychologischen Phänomene nicht beobachtbar (und daher auch nicht existent) seien, die sich nicht mit den bekannten - vor allem physikalischen - Meßverfahren messen ließen. Für viele Psychologen bedeutete dies eine Blockierung jeder Forschung auf zentralen Gebieten der Psychologie. Jene, die nicht ganz so weit gehen wollten wie Watson, sagen, daß psychologische Dispositionen nicht *direkt* beobachtbar, aber anhand von Indikatoren und bestimmten Zusatzannahmen erschließbar seien. Man spricht hier von "hypothetischen" oder "latenten Konstrukten". Aber diese Position bewegt sich wegen der unbeweisbaren Annahmen, die sie voraussetzen muß, auf höchst unsicheren, wissenschaftlichen Grundlagen.<sup>4</sup>

Die methodologische Position unserer Untersuchung ist, daß psychische Dispositionen wie moralische Urteilsfähigkeit und moralische Einstellungen als manifeste Eigenschaften des Verhaltens begriffen und gemessen werden können (s. Kapitel 2 und 4). Nach dieser, als *methodologischer* Behaviorismus

4 Für eine ausführlichere Darlegung meiner Kritik siehe Lind, 1985a, und Kap. 3.



bezeichneten Position, muß keineswegs angenommen werden, daß moralische Prinzipien bloß latente Konstrukte sind, die nur unterstellt ("hypothetische Konstrukte") oder intuitiv erfaßt, aber nicht empirisch-experimentell nachgewiesen werden können. Es ist auch nicht notwendig, sie bloß als eine willkürliche Beschreibung bestimmter Verhaltensakte und daher als weitgehend selbstreferentiell aufzufassen (im Sinne der Aussage: "Die Person A ißt, also nenne ich sie hungrig").

Im Zentrum der vorliegenden Untersuchung steht der Begriff der "moralischen Urteilsfähigkeit" und seine Operationalisierung als Struktur des Urteilsverhaltens in moralisch relevanten Situationen. Nach Kohlberg (1964, S. 425)

*ist moralische Urteilsfähigkeit "das Vermögen, moralische Entscheidungen und Urteile zu treffen, die moralisch sind (also auf moralischen Prinzipien gründen), und in Übereinstimmung mit diesen Urteilen zu handeln."*

Diese Definition steht in Kontrast zu zwei traditionsreichen Definitionen, von denen zu zeigen ist, daß sie weniger adäquat sind. Die eine definiert moralisches Verhalten als einen Akt, der – äußerlich – mit sozialen Konventionen und Normen übereinstimmt. Demnach wäre derjenige moralisch zu nennen, der eine als "moralisch" definierte Handlung durchführt oder eine "unmoralische" unterläßt. Diese Definition von Moral ist vor allem aus zwei Gründe inadäquat: 1. Eine isoliert betrachtete Handlung läßt fast nie einen eindeutigen Rückschluß auf die Moralität des Handelnden zu. Eine Handlung kann von anderen als "moralisch" betrachtet werden, ohne daß die Person moralisch sein wollte, oder sie kann allein dem Zweck dienen, von anderen als moralisch eingeschätzt zu werden, damit ein dahinterliegender - vielleicht unmoralischer - Zweck erfüllt wird. Jemand verschenkt zum Beispiel seinen Mantel an einen Bettler, weil er ihm nicht mehr gefällt (=nicht-moralisches Motiv), und sorgt dafür, daß die Presse über das Geschenk berichtet (=übergeordneter Zweck). 2. Die Definition von "moralisch" als die Übereinstimmung einer Handlung mit den sozialen Normen hat so viele Bedeutungen wie es sozio-kulturelle Normensysteme gibt. Ein und dieselbe Handlung müßte einmal als moralisch, ein andermal als unmoralisch beurteilt werden, je nachdem welche soziale Norm gerade als Kriterium herangezogen wird. Zum Beispiel das Verhalten: Ein Schüler läßt einen

anderen während einer Klassenarbeit abschreiben. Aus der Sicht des Lehrers ist das ein Verstoß gegen die Regeln der Schule, aus der Sicht des Schülers aber eine Erfüllung des Hilfegebots unter Freunden. Der Definition von Moral als einfache Normkonformität fehlen also zwei wichtige Bestimmungsstücke, nämlich daß der Akt auf moralischen Prinzipien des Individuums beruht und daß diese Prinzipien universalisierbar sind.

Die zweite, linguistische Definition von Moral als *Begründungen von Verhalten mit moralischen Begriffen* (= moralische Aussagen) unterstellt einen Bezug zum Handeln, kann ihn aber nicht belegen. So läßt sich zum Beispiel bei ihr nicht entscheiden, ob Aussagen mit moralischem Inhalt eine bestimmende Funktion für das Verhalten einer Person haben, oder ob sie nur Ausdruck einer intellektuellen Übung oder nachträglicher Rechtfertigung ("Rationalisierung") von Meinungen und Entscheidungen sind.<sup>5</sup> Moralische Aussagen bieten zwar einen Zugang zu den moralischen Dispositionen eines Menschen, aber für sich betrachtet ermöglichen sie keine gültige Messung der moralischen Urteilsfähigkeit. Erst der psycho-soziale Kontext, in welchem moralische Argumente benutzt werden, bestimmt ihre psychologische Realität. Ein Schauspieler, beispielsweise, der moralische Sätze rezitiert, wird nicht deshalb schon urteilsfähig genannt werden. Wenn sich aber eine Person in einer wichtigen Sache durch ein moralisches Argument überzeugen läßt, obwohl es ihrer eigenen Meinung entgegengesetzt ist, sind wir bereit zu unterstellen, daß bei ihr moralische Motive vorhanden sind und daß sie fähig ist, diese Motive auf die Steuerung ihres eigenen Tuns anzuwenden.

Die Probleme beider Ansätze werden in der Kohlbergschen Definition von moralischer Urteilsfähigkeit vermieden. Zum einen wird ein Handeln nur soweit "moralisch" genannt, wie es nachweislich durch moralische Prinzipien

5 Kohlberg (1984) versuchte, mit Hilfe des Konstrukts des *Verantwortungsurteils* zwischen moralischen Aussagen und Handeln eine empirische Verbindung herzustellen. In einer exzellenten Untersuchung metaethischer Positionen der Psychologie hat Thomas Wren (1990) diesen Versuch verworfen, weil er einen infiniten Regreß der Urteile impliziert (wie kann die Brücke zwischen *Verantwortungsurteilen* und Handeln geschlagen werden?). Wren postuliert dagegen eine intrinsische Verbindung zwischen jedem *ernsthaft* geäußerten moralischen Argument und dem Handeln einer Person. Aber auch diese Hypothese der Ernsthaftigkeit kann im Einzelfall nicht geprüft werden und läßt somit das Problem ungelöst.

determiniert ist. Zum anderen werden moralische Prinzipien (bzw. ihre verbale Repräsentation in Form von Argumenten) nur soweit als Ausdruck dieser Fähigkeit angesehen, wie sie sich verhaltenswirksam erweisen. Moralische Urteilsfähigkeit läßt sich weder allein durch das Verhalten, noch allein durch das ihm zugrundeliegende Denken definieren, sondern immer nur durch beides. Nach der Kohlberg-Definition, auf die auch wir uns im großen und ganzen stützen, macht streng genommen die Frage keinen Sinn, wie Urteilen und Handeln *empirisch* zusammenhängen. Sie lassen sich unabhängig voneinander gar nicht definieren und messen.<sup>6</sup>

Dies zeigt, wie eng inhaltliche und methodologische Fragen miteinander verbunden sind. Ohne die Möglichkeit, das Konstrukt "moralische Urteilsfähigkeit" so zu messen, wie es theoretisch konzipiert ist, kann die Frage nicht empirisch entschieden werden, welche der Thesen zur Moralentwicklung zutreffend ist. Lange Zeit wurde Moral – in Übereinstimmung mit der Sozialisationsthese – nur als Einhaltung oder Abweichung von sozialen Normen gemessen (moralisch kompetent ist, wer die Normen kennt und einhält).<sup>7</sup> Piaget und Kohlberg haben als erste erkannt, daß der Kern der Moralität eine über das Erlernen sozial anerkannter Normen hinausgehende *kognitive* Fähigkeit darstellt und diese Erkenntnis in geeignete Meßinstrumente zu übersetzen versucht.

Die größte Aufmerksamkeit hat dabei Kohlbergs *Moral Judgment Interview* (MJI) erfahren, das in Kapitel 3 näher dargestellt wird. Wie dort gezeigt

- 6 Davon zu unterscheiden ist die im Alltag interessierende Frage, ob sich aus der Kenntnis einer Verhaltensdisposition einer bestimmten Person (z.B. Herr X ist kein Betrüger) ihre Entscheidung in einer bestimmten moralischen Konfliktsituation vorhersagen läßt (Herr X nimmt kein Geld aus der Kasse, wenn er allein im Laden ist). Dies ist aber kein Urteil-Handlungs-, sondern ein Handlung-Handlungs-Problem, das sich prinzipiell empirisch untersuchen läßt. Praktisch aber sind aus ethischen und anderen Gründen nicht alle Handlungen der Forschung zugänglich. Nur als Literaturgestalt konnte Dostojewskijs *Raskolnikow* einen Menschen umbringen, um die Wirkung dieser Tat auf seine moralischen Gefühle studieren zu können. Wie eng die Grenzen des Experimentierens auf dem Gebiet der Moral sein können, zeigt die Diskussion um das *Milgram-Experiment* (Milgram 1974; für einen Überblick siehe Meeus & Raajmakers (1989).
- 7 Siehe die gründliche Kritik von Pittel & Mendelsohn (1966) an dieser Art der operationalen Definition und Messung von Moralität.

wird, erfüllt das MJI in seiner neuesten Revision voll die Kriterien der sogenannten klassischen Testtheorie, was ein wichtiges Ziel dieser Revision war. Aber dieser "Erfolg" hat sich, so soll gezeigt werden, negativ auf die theoretische Validität dieses Meßinstruments ausgewirkt. Zudem führen bestimmte Anweisungen des Handbuches (Colby, Kohlberg et al. 1987) zur Bevorzugung der Invarianztheorie, was einen fairen Theorievergleich erschwert. Diese Verzerrung zugunsten der Reifungstheorie wird zudem verstärkt durch die weit verbreitete Tendenz in der *Moral judgment*-Forschung, das Bildungsniveau als Bedingungsfaktor zu übersehen und fast nur Jugendliche zu untersuchen, die sich in der Ausbildung befinden, bei denen also altersmäßige Reifung und Bildungserfahrungen zusammenfallen.

Zur Behebung dieser Probleme wird in dieser Untersuchung ein neuer Meßansatz angewandt, der *Moralisches Urteil-Test* (MUT), der eine unverzerrte Messung von Regressionen zuläßt (Kapitel 4). Der MUT orientiert sich zwar eng an Piagets und Kohlbergs *Theorie* der Moral, insbesondere an der Struktureigenschaft des moralischen Urteilverhaltens, aber nicht an deren Methoden. Die Operationalisierung der Struktur basiert vor allem auf den Methoden der psychologischen Einstellungsforschung (Likert-Skalen), der experimentellen Psychologie (multifaktorielle Designs), der *Judgment*-Skalierung und der – in Vergessenheit geratenen – *diakritischen Methode* zur Überprüfung von Dispositionshypothesen von Egon Brunswik.<sup>8</sup> Zum anderen wurden neben Jugendlichen, die lange in der Ausbildung verblieben, auch solche untersucht, die ihre Ausbildung früh beendeten, so daß der Einfluß von Alter und Bildungsniveau auf die Moralentwicklung – zumindest partiell – unabhängig voneinander studiert werden kann.

### **1.3 Kritik des Kohlberg-Ansatzes**

Aufgrund des Kohlberg-Ansatzes wurden zwei wichtige Entdeckungen möglich: Erstens, fast alle Menschen anerkennen unabhängig von Geschlecht, Schicht oder Kultur moralische Prinzipien wie Gerechtigkeit und Hilfsbereitschaft. Nahezu jeder bevorzugt dabei – im Sinne der Kohlberg-Stufen – höhere Formen der moralischen Argumentation gegenüber niedrigeren (Rest

8 Zur *Judgment*-Skalierung s. Torgerson 1967/1958; zur diakritischen Methode s. Brunswik 1955; auch Kapitel 4.

1973; Lind 1985a). Zweitens, die Fähigkeit, moralische Konzepte im eigenen Verhalten konsistent und differenziert anzuwenden, also das, was wir als moralische Urteilsfähigkeit bezeichnen, ist in jungen Jahren noch wenig ausgebildet und bedarf der Entwicklung. Kinder wenden moralische Prinzipien oft noch inkonsistent und undifferenziert an. Wenn ihr Bewußtsein mit der Verfolgung *eines* Ziels beschäftigt ist, haben andere Ziele kaum eine Chance. Auf den unteren Stufen der individuellen Moralentwicklung lassen sich Menschen noch selten von moralischen Argumenten beeinflussen, insbesondere wenn sie gerade versuchen, ihre Meinung durchzusetzen. Erst im Verlauf der Entwicklung lernen sie, das Engagement für eine Sache mit einem integrierten und differenzierten moralischen Urteil zu verbinden. Erst dann können sie mehrere Ziele, moralische und nicht-moralische bei einer Entscheidung gleichzeitig berücksichtigen und in ihrem Selbst integrieren.

Die vorliegende Untersuchung zeigt, daß der Kohlberg-Ansatz aber stark revisionsbedürftig ist bezüglich des Verhältnisses von Affekt und Kognition im moralischen Verhalten, der Definition von "Struktur" und, worauf bereits hingewiesen wurde, der Annahme der invarianten Sequenz: a) *Moralische Affekte und Kognitionen dürfen nicht als separate Verhaltenskategorien, sondern müssen als Aspekte ein und desselben Verhaltens verstanden und gemessen werden.* In seiner ersten Arbeit hierzu, definierte Kohlberg (1958, S. 89) dieses Verhältnis ganz unterschiedlich: (A) als eine untrennbare Einheit, (B) als unterscheidbare, aber parallele Aspekte des moralischen Verhaltens, © als Stufe und Ebene der Moralentwicklung und (D) als moralische und logische Urteilsfähigkeit. Die Gleichsetzung von Stufen mit Affekten (Einstellungen) und Ebenen mit Kognitionen hat Kohlberg später zu Recht fallengelassen. Die Gleichsetzung des kognitiven Aspekts der Moral mit a-moralischen, logischen Denkfähigkeiten war wohl eine Verlegenheitslösung, da ihm keine Methoden zur Verfügung standen, um kognitive Aspekte der Moral unabhängig von den affektiven zu konzeptualisieren und messen zu können. Es verbleiben die Definitionen (A) und (B). In etwas modifizierter Fassung bilden sie die Grundlage unseres "Zwei-Aspekte-Modells" der moralisch-kognitiven Entwicklung und unseres Meßansatzes (s. Kapitel 2 und 4). Mit dem Zwei-Aspekte-Modell lassen sich zentrale Widersprüche des Stufen- und Struktur-Begriffs bei Kohlberg auflösen. Kohlberg hat in seinen Stufen lange Zeit eine Verlängerung der Piaget-Phasen gesehen. Die Gleichsetzung von Stufe 1 und 2 mit Piagets Heteronomie und Autonomie war jedoch nicht

haltbar, ebensowenig wie die Gleichsetzung von moralischer Autonomie mit Postkonventionalität. Mit der Identifikation von Unterstufen (*A* und *B*) hat Kohlberg später die Piaget-Phasen unabhängig von seinen Stufen zu definieren versucht. Dieser Versuch hat aber die Probleme bei der Operationalisierung von Struktur und struktureller Ganzheit nur noch deutlicher gemacht. Wie lassen sich Unterstufen vorstellen, wenn Stufen eine strukturelle Ganzheit bilden?

b) *Die Struktur des moralischen Verhaltens darf nicht als ein bestimmter Verhaltensinhalt (auf die die Strukturdefinition im Auswertungshandbuch paßt) verstanden werden, sondern muß relational definiert und gemessen werden.* Struktur bedeutet üblicherweise die Menge der *Relationen*, die zwischen einzelnen Elementen bestehen. Für die Zuerkennung moralischer Motive ist zum Beispiel die relationale Information wichtig, wie *konsistent* eine Person ein moralisches Argument verwendet. *Konsistenz* kann man nicht messen, wenn man nur einzelne Akte einer Person isoliert für sich betrachtet, wie es die klassische Psychometrie tut, sondern nur, indem relationale Merkmale von Antwort*mustern* analysiert werden. In Kohlbergs Interviewmethode ist Struktur nicht relational, sondern *formal* definiert. Obwohl Kohlberg und seine Kollegen diese Definition von Struktur ständig zu präzisieren versucht haben, ist sie bis heute umstritten (Eckensberger & Reinshagen 1980; Lind 1985b; Keller 1990). Bemerkenswerterweise definiert Kohlberg in seinen theoretischen Arbeiten Struktur *relational*, nämlich als das manifeste *Muster* des individuellen Urteilsverhaltens. Eine relationale Definition von Struktur wird auch durch das Postulat der "strukturellen Ganzheit" nahegelegt, das Kohlberg durch die hohe Antwortkonsistenz seiner Befragten bestätigt sieht. Aber Kohlberg und seine Kollegen haben sich für eine a-strukturelle Methodologie (die sog. klassische Testtheorie) als Grundlage ihrer Forschung entschieden, was die Erhebung relationaler Informationen, wie Konsistenz, ausschließt (Kapitel 3).

c) Die Moralentwicklung ist gekennzeichnet durch einen relativ langsamen Aufbau von Fähigkeiten (statt einer schnellen Änderung von Einstellungen) infolge von Bildungsprozessen und einem mindestens ebenso langsamen Abbau dieser Fähigkeiten nach Wegfall von Bildungs- und Übungsmöglichkeiten. Moralentwicklung ist also nicht, wie Kohlberg annimmt, gekennzeichnet durch eine invariante Abfolge von Stufen; *moralische Urteilsfähigkeit entwickelt sich zurück, wenn sie nicht durch diskursive*

*Prozesse und pädagogische Interventionen stabilisiert und gefördert wird* (s. Kapitel 5 und 11).

#### **1.4 Die sozialpsychologische Perspektive der Moral judgment-Forschung**

Die Sozialpsychologie ist, so stellt Kumpf fest, "in den USA und auch hierzulande zusehends a-sozialer geworden, in dem Sinne, daß sie sich ganz dem Studium von Individuen zu- und von sozialen Phänomenen abwandte" (Kumpf 1990, S. 123). Dieser Vorwurf wird auch der kognitiven Entwicklungstheorie gemacht (s. Schreiner 1992), was aber nur zum Teil berechtigt ist. Schon früh hat Kohlberg (1969) versucht, individuelle Eigenschaften auf der einen Seite und Merkmale sozialer Institutionen auf der anderen, sowie ihr Aufeinandertreffen und ihr Zusammenwirken theoretisch miteinander in Beziehung zu setzen. Weniges davon konnte aber in empirische Forschung übersetzt werden. Die vorliegende Untersuchung greift daher die sozialpsychologische Dimension verstärkt auf, indem Kommunikationsprozesse, Phänomene der sozialen Wahrnehmung, demokratische Interaktionen und Prozesse der sozialen Wahrnehmung – Stereotypisierung und pluralistische Ignoranz – in die Analyse mit einbezogen werden (Kapitel 8 und 10). Darüber hinaus knüpft sie, wie schon erwähnt, am klassischen Einstellungskonzept der Sozialpsychologie an bzw. baut dieses strukturell aus. Das Einstellungskonzept hat, wie Scott (1968, S. 208) zu Recht bemerkte, den unschätzbaren Vorteil, daß seine präzise Definition und Empiriebezogenheit seine Unangemessenheiten offenlegt. Vor Kohlberg wurde lange Zeit fast ausschließlich dieses Konzept verwendet, um die Moralität von Personen zu analysieren.<sup>9</sup> Einstellungsskalen liefern unmittelbares verbales Verhaltensmaterial. Sie sind daher nicht mit den Problemen (Halo-Effekt, Anker- und Referenzreiz-Problemen) der Selbstbeschreibungen oder Selbstdiagnosen in Persönlichkeitstests behaftet.<sup>10</sup> Das Manko des klassi

9 May & Hartshorne 1926; Hartshorne & May 1928; Krathwohl et al. 1964; Shaw & Wright 1967; zusammenfassend Pittel & Mendelsohn 1966.

10 Es liegt nahe, systematische Tendenzen von Menschen bei der Beurteilung anderer (R. Cohen 1969) auch dort zu vermuten, wo es um die Beurteilung der  
(Fortsetzung...)

schen Einstellungskonzepts liegt darin, daß es keine Möglichkeit bietet, *strukturelle* Unterschiede festzustellen, so daß es zum Beispiel nicht möglich ist, undifferenzierte Einstellungen von neutralen Einstellungen zu unterscheiden (Scott 1968; Converse 1970). Durch die strukturelle Erweiterung des Einstellungskonzept im Rahmen unseres Zwei-Aspekte-Modells ist dieses Problem weitgehend gelöst (Kapitel 2 und 4; Lind 1982). Das um das Strukturmerkmal erweiterte Einstellungskonzept ist zur Zeit offensichtlich die einzige Möglichkeit, für uns wichtige Thesen wie die der Regression, der Kulturspezifität oder der affektiv-kognitiven Parallelität adäquat zu überprüfen. Darüber hinaus erfüllt es die Ockham-Russelsche Forderung nach Intersubjektivität und nach theoretischer Transparenz und Sparsamkeit<sup>11</sup>, vor allem im Vergleich zu sogenannten qualitativen Meßansätzen, die sich oft auf *ad hoc* formulierte, nur schwer nachvollziehbare Annahmen stützen.<sup>12</sup>

10(...Fortsetzung)  
eigenen Persönlichkeit geht.

- 11 Die Forderung des Wilhelm von Ockham "entia non sunt multiplicanda, sine necessitate" wird auch als "entia non sunt multiplicanda praeter necessitatem" überliefert. Frei übersetzt: Man solle die Zahl der Konstrukte nicht ohne Not vermehren. Der schneidenden Kritik wegen, mit der Ockham (1290-1350) diesen Satz gegen die scholastische Unart gebraucht hat, jeder in einer Theorie auftretenden Schwierigkeit mit der Erfindung immer neuer Konstrukte zu begegnen, ist er in der Philosophie auch als "Ockhams Rasiermesser" bekannt. Die moderne Version dieser Forderung stammt von B. Russel (1972): "Wenn in einer Wissenschaft alles interpretiert werden kann, ohne diese oder jene hypothetische Sache anzunehmen, dann gibt es keinen Grund sie anzunehmen" (S. 472; soweit nicht anders vermerkt sind alle fremdsprachigen Zitate in dieser Arbeit vom Verfasser übersetzt).
- 12 Kohlberg hat seine Skepsis gegen das Einstellungskonzept erst in jüngerer Zeit revidiert und anerkannt, daß mit strukturellen Einstellungstests wie dem *Moralisches Urteil-Test* (MUT) wesentliche Verbesserungen erreicht werden können: Mit dem MUT, so Kohlberg (1985), "ist ein Auswertungsalgorithmus möglich, mit dem sich jedem Individuum ein reiner Strukturwert zumessen läßt" (S. xvii).



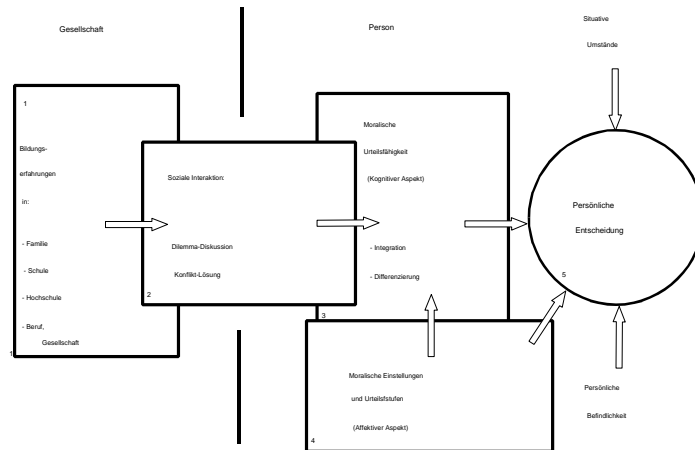


Abb. 1-1 Ein Prozeßmodell

### 1.5 Ein Prozeßmodell der Veränderung der moralischen Urteilsfähigkeit

Der Aufbau der hier vorgelegten Arbeit orientiert sich an einem allgemeinen Prozeßmodell der Veränderung der moralischen Urteilsfähigkeit. Dieses Prozeßmodell spiegelt in groben Zügen die Grundannahmen wider, die wir hier auf den Prüfstand stellen wollen, und bildet gleichzeitig ein Resümee unserer Untersuchung (Abb. 1-1). Alle unsere Analysen gründen auf offen *beobachtbarem Verhalten* von Menschen und nicht auf prinzipiell unverifizierbaren ("hypothetischen") Konstrukten oder unbeweisbaren Annahmen über Motive von Menschen. Wo wir mentale Begriffe bei der Beschreibung des moralischen Verhaltens verwenden, versuchen wir diese so präzise wie möglich als Eigenschaften des *Verhaltens* eines Individuums operational zu definieren, statt allein auf den unsicheren Grund unserer Intuition zu vertrauen. Aber wir verbinden mit unseren Untersuchungen an keiner Stelle den Anspruch, einen ganz bestimmten Verhaltensakt eines Menschen *vorherzusagen*, wie es schon so viele vergeblich versucht haben. Dieser Verzicht resultiert sowohl aus einer ethischen Selbstbeschränkung wie auch aus der Einsicht in die prinzi-

pielle Unmöglichkeit eines solchen Unterfangens. Wir können nur davon ausgehen, daß die moralische Urteilsfähigkeit ebenso wie die moralischen Einstellungen einer Person einen wichtigen Einfluß auf ihre persönliche Entscheidung haben. Aber wir müssen ebenso davon ausgehen, daß diese Entscheidung, soll sie angemessen und gerecht sein, ebenso sehr von – unvorhersehbaren – situativen Umständen und Befindlichkeiten abhängt. Die vorliegende Untersuchung beschränkt sich daher auf die moralische Urteilsfähigkeit und die Bedingungen ihrer Entwicklung. Dies schließt nicht aus, daß moralische Urteilsfähigkeit eine große Bedeutung für die Güte individueller Entscheidungen hat. Es spricht sogar vieles für diese Vermutung. Aber es soll dem Leser überlassen bleiben, ob er die vorliegende Evidenz dafür für ausreichend hält. Unser Modell zeigt daher die persönliche Entscheidung des Einzelnen als Leerstelle.

Die Kapitel dieser Arbeit sind den Komponenten unseres Prozeßmodells zuordenbar; die nachfolgenden Kapitel beginnen mit dem zentralen Aspekt der moralischen Urteilsfähigkeit (Abb. 1-1, Kasten 3). Kapitel 5 untersucht anhand des Regressionsphänomens die Reifungsthese. Die Kapitel 6 und 7 richten sich auf die Hypothesen, die aus der Sozialisationsthese folgen. Die folgenden Kapitel enthalten Teiluntersuchungen zu verschiedenen Aspekten und "Mechanismen" der Veränderung des moralischen Urteils in sozialen Interaktionen im Kontext von Bildungsprozessen, und zwar zur Möglichkeit der effektiven Veränderung des "moralischen Klimas" eines institutionellen Kontextes (Kapitel 8), zur Wirksamkeit sozialer Interventionen, die sich gezielt auf eine Stimulation der kognitiven Struktur des moralischen Urteilsverhaltens richten (Kapitel 9) und zur Genauigkeit der Wahrnehmung von "moralischen Tatsachen" im institutionellen Kontext Schule (Kapitel 10). Im Kapitel 11 wird abschließend die Frage untersucht, welchen Einfluß Bildungserfahrungen im Vergleich zu anderen Faktoren wie Sozialisation, Alter und Selektion auf die moralisch-kognitive Entwicklung haben (Abb. 1-1, Kasten 1).

## **1.6 Technische Angaben**

### *1.6.1 Die Datenbasis dieser Untersuchung*

Die Hypothesen dieser Untersuchung werden an verschiedenen Arten von empirischen und experimentellen Daten überprüft. Der überwiegende Teil der

Daten, über die hier berichtet wird, stammt aus Untersuchungen im Rahmen von größeren Forschungsprojekten oder eigenen Projekten. Die Hypothesen wurden aber auch an Ergebnissen anderer Untersuchungen überprüft, wofür diese in vielen Fällen neu berechnet oder durch geeignete Umrechnungen vergleichbar gemacht wurden. Für diese Berechnungen wurden meist die Originalberichte oder manchmal auch die Rohdaten herangezogen. In einigen Fällen wurden die Ergebnisse nur in eine graphische Darstellung übersetzt, um besser beurteilen zu können, welche These sie stützen. Auf weitere Einzelheiten und die Quellen wird jeweils verwiesen. Ein wesentlicher Teil der Datengrundlage zu dieser Arbeit stammt aus der *international vergleichenden Längsschnittstudie* der Konstanzer Forschungsgruppe Hochschulsozialisation.<sup>13</sup> Hieraus wurden, je nach Fragestellung, verschiedene Teilgruppen analysiert (Kapitel 5, 7 und 11). Die Definition dieser Teilgruppen wird in den entsprechenden Kapiteln angegeben. Als weitere Datengrundlage für diese Untersuchung dienen ein *Laborexperiment* zur Frage der Simulation von moralischer Urteilsfähigkeit (Kapitel 6) und ein *Feldexperiment* zur Veränderung der moralischen Atmosphäre in der Schule (Kapitel 8). Letzteres wurde in Zusammenarbeit mit Fritz Oser (Fribourg) und dem Landesinstitut für Schule und Weiterbildung des Landes Nordrhein-Westfalen geplant und durchgeführt. Die Daten der Repräsentativ-Untersuchung über die moralische Urteilsfähigkeit bei Jugendlichen und jungen Erwachsenen (Kapitel 5) entstammen einer 1990 durchgeführten, unveröffentlichten Berufsbildungsstudie von EMNID, Bielefeld.

13 Mitglieder der Forschungsgruppe waren neben dem Autor Tino Bargel, Barbara Dippelhofer-Stiem, Gerhild Framhein, Hansgert Peisert (Projektleiter), Johann-Ulrich Sandberger und Hans Gerhard Walter. Auf internationaler Ebene war die Studie Teil der vom Europäischen Koordinationszentrum für Sozialwissenschaftliche Forschung in Wien koordinierten, international vergleichenden Längsschnittstudie "Diplomés de l'Université: leur *Formation* et leur *Conception de la Vie*" (kurz: FORM) unter der Leitung von Hansgert Peisert und Władisław Markiewicz. Vgl. Sandberger et al. 1982, Framhein & Langer 1984, Dippelhofer-Stiem & Lind 1987.

### 1.6.2 Zur statistischen Bewertung von Zusammenhängen

Die vorliegende Untersuchung ist relativ stark von methodologischen und forschungsmethodischen Überlegungen geprägt. Dies erwies sich als notwendig, da in diese Methoden unser ganzes Vorverständnis eingeschmolzen ist, mit dem wir uns dem Gegenstand unserer Untersuchung nähern. Sie sollen daher hier näher erläutert werden (s. Lind 1985a).

Wenn wir unsere Forschung auf inadäquate Methoden, d.h. auf falsches Vorverständnis gründen, kann sie nicht viel erbringen. Unglücklicherweise glauben viele Sozialwissenschaftler, daß bestimmte, historisch gewachsene Methoden von absoluter Gültigkeit, und damit *der* Ausweis für Wissenschaftlichkeit und die Basis fachlicher Zugehörigkeit seien - womit sie jede kreative Forschung, die immer auch auf der Entwicklung neuer Methoden beruht, zum Erliegen bringen. Die wirkliche Basis kollegialer Zusammenarbeit ist das Interesse an einem gemeinsamen Thema, der richtige Ausweis für Wissenschaftlichkeit das Ideal, möglichst informative oder, wie Popper (1968) sagt, "riskante" Hypothesen aufzustellen, und sie einem fairen und intersubjektiv nachvollziehbaren Test durch Beobachtungen zu unterziehen. Hier wird also nicht der methodologische Standpunkt des "anything goes" vertreten. Im Gegenteil, es wird verlangt, daß die verwendeten Methoden nicht nur eine gehaltvolle und objektive Prüfung von Theorien garantieren, sondern auch, daß sie eine valide Operationalisierung der Fragestellung sind.

Im Rahmen unserer Untersuchung, in deren Mittelpunkt die Prüfung dreier Theorien über die Natur und die Bedingungen der Moralentwicklung steht, interessiert vor allem die Frage, ob und in welchem Maße diese Theorien bzw. die daraus ableitbaren Hypothesen im Einklang mit den Beobachtungen stehen. Diese Fragestellung unterscheidet sehr von solchen Studien, die mit dem Ziel durchgeführt werden, an einer kleinen Gruppe Daten zu gewinnen, die - mittels der Inferenzstatistik - auf größere Gruppe verallgemeinert werden können. Daher werden hier, wie der Statistiker Kruskal (1958, S. 815) vorschlägt, Untersuchungsgruppen nicht als Stichproben, sondern als Populationen behandelt. Maße der statistischen Signifikanz wie  $t$ ,  $F$  oder  $\chi^2$  sichern nicht generell, wie ihr Namen suggerieren mag, die "Bedeutsamkeit" oder "Verallgemeinerungsfähigkeit" eines

Befundes.<sup>14</sup> Zudem gibt die Höhe der statistischen Signifikanz keinen Aufschluß über die Güte einer Theorie bzw. einer Vorhersage, da sie vor allem von der Größe der Stichprobe – also einer frei wählbaren Größe – abhängig ist. Statt, wie mancherorts vorgeschlagen, diesen Nachteil durch komplizierte Maßnahmen zur Sicherung der "Teststärke" (power) auszugleichen, erscheint es naheliegender, direkt Maße für die Effektstärke zu berechnen (Cohen 1988). Diese Maße stellen mathematisch gesehen nichts anderes dar als eine Standardisierung der Signifikanzmaße auf die Stichprobengröße. Es handelt sich hier im wesentlichen um Kennziffern wie den Korrelationskoeffizienten  $r$ , Varianzanteile oder  $d$ -Werte (Glass et al. 1978; Cohen 1988). Wir werden uns in dieser Arbeit in der Hauptsache auf  $r$  als Maß für die Stärke eines Effekts bzw. die Größe eines Zusammenhangs oder einer Differenz stützen (s. Kapitel 9).

Aber auch Effektstärkemaße haben Nachteile. Sie sind, entgegen der oft benutzten Kennzeichnung als "Maße der praktischen Relevanz", nicht immer geeignet, um praktisch relevante Aussagen über die Größe eines Effekts zu formulieren. Sie sind nicht unabhängig von zufälligen, für die Aussage irrelevanten Charakteristika der jeweiligen Untersuchungsgruppe. Ist zum Beispiel die Varianz einer Variablen, deren Effekt untersucht werden soll, in einem Sample zufällig sehr gering oder null, dann kann sich kein Effekt zeigen, auch wenn er vorhanden wäre. Eine naheliegende Möglichkeit der Lösung dieses Problems besteht darin, die absoluten Meßwertunterschiede, die in einer Studie registriert werden, *direkt* als Maß für die Effektstärke zu nehmen. Wenn zwei Meßwerte, z.B. 4,3 und 5,1, vorliegen, so ist der Effekt 0,8. Hier wird man aber zurecht einwenden, daß die Zahl 0,8 nur sinnvoll interpretiert werden kann, wenn man weiß, wie groß der Wertebereich ist, den die Skala zuläßt. Bei einer Skala von 4 bis 6 stellt eine Differenz von 0,8 einen relativ großen Wert dar; bei einer Skala von -100 bis +100 aber einen sehr kleinen Wert. Die Bedeutung absoluter Differenzen hängt also ganz entscheidend von den Werten ab, die bei einer bestimmten Skala aufgrund ihrer

14 Diese Maße geben nur an, inwieweit ein Befund repräsentativ für die Grundgesamtheit von Fällen ist, von der eine Stichprobe ausgewählt wurde. Die Stichprobenbefunde sind nur dann repräsentativ, wenn die Gesamtheit genau definiert und die Auswahl streng nach Zufall geschehen ist (Hays 1963).

Konstruktion theoretisch erreicht werden können. Einige Autoren wie beispielsweise Fend und seine Mitarbeiter (Fend et al. 1976) haben daher alle Skalen, die sie in einer Untersuchung verwendet haben, auf eine gleiche Länge (0 bis 20) umgerechnet und so eine Vergleichbarkeit hergestellt. Diese Idee in etwas abgewandelter Form aufgreifend, will ich zwei Schritte zur Erzielung einer standardisierten (von irrelevanten Gegebenheiten unabhängigen) Effektstärke vorschlagen: Erstens die Berechnung einer *relativen Differenz* (die absolute Differenz dividiert durch die absolute Skalenbreite) und zweitens eine Übereinkunft darüber, wie groß eine relative Differenz sein muß, um als "bedeutsam" oder "sehr bedeutsam" bezeichnet zu werden. Eine solche Übereinkunft wurde, soweit ich weiß, bislang von niemanden explizit getroffen, aber sie ist in der empirischen Forschung faktisch vorhanden. Die Durchsicht sehr unterschiedlicher sozialwissenschaftlicher Studien zeigt, daß zwar die Spannweite der Differenzwerte, die in der Literatur als *sehr* oder *hoch bedeutsam* bezeichnet werden, groß ist; wir fanden die Bezeichnung "hoch signifikant" für relative Differenzen von 1,5 bis 15 Prozentpunkte. Aber etwa drei Viertel aller Differenzwerte, die so bewertet wurden, lagen bei 8 bis 10 Prozentpunkten. Das heißt, in der Regel werden Differenzwerte, die eine Größe von ca. 8 bis 10 Prozent der theoretisch möglichen Skalenbreite betragen, als "hoch signifikant" oder "sehr bedeutsam" bezeichnet. Folglich werde ich in dieser Untersuchung den Terminus "sehr bedeutsam" für Effekte zu verwenden, die gleich oder größer 10 Prozent der Skala sind, mit der sie gemessen wurden. Eine relative Differenz von 5 Prozent soll bloß als "bedeutsam" bezeichnet werden. Mit diesen beiden Eckwerten haben wir zwei "praktikable" Kriterien für die Interpretation von Differenzwerten, wo die isolierte Bewertung von Effekten angesagt ist. In den Fällen, in denen ein direkter Vergleich zweier oder mehrerer Vorhersagen mit den Beobachtungen möglich ist, sind solche Interpretationshilfen ohnehin unnötig.

#### Anmerkungen

a. Brief von Professor Rossi im Jahr 1984.