

**Materialien aus der Bildungsforschung Nr. 55**

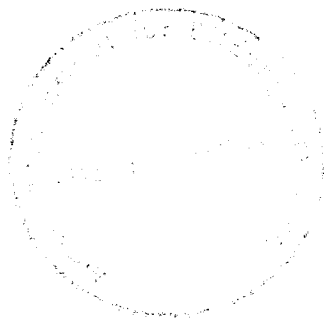
**Siegfried Reuss und Günter Becker**

**EVALUATION DES ANSATZES VON  
LAWRENCE KOHLBERG ZUR ENTWICKLUNG  
UND MESSUNG MORALISCHEN URTEILENS**

**Immanente Kritik und Weiterentwicklung**

**Max-Planck-Institut für Bildungsforschung  
Berlin 1996**

**GW ISSN 0173-3842  
ISBN 3-87985-048-8**



E 96/591 +3

Materialien aus der Bildungsforschung

(3. Ex.)

In dieser Reihe veröffentlicht das Institut für Bildungsforschung technische Berichte und andere Materialien aus der Forschung, die in der Regel keine abgeschlossenen Forschungsberichte sind, aber dem jeweils interessierten Fachpublikum zugänglich gemacht werden sollen.

Bestellungen werden erbeten an die Verwaltung des Instituts bei gleichzeitiger Überweisung von DM 13,- (einschließlich 7% Mehrwertsteuer) auf das Konto Nr. 0910005885 der Berliner Sparkasse, BLZ 100 500 00.

Nachdruck, auch auszugsweise, ist nur mit Zustimmung des Instituts gestattet.

©1996 Max-Planck-Institut für Bildungsforschung, Lentzeallee 94, D-14195 Berlin.

GW ISSN 0173-3842  
ISBN 3-87985-048-8

# Inhalt

Zusammenfassung/Summary .....	2
Vorbemerkung .....	3
1. Kohlbergrekonstruktion im Schnittpunkt von drei Analyseebenen .....	6
1.1 Kohlbergs Interview- und Auswertungsverfahren .....	6
1.2 Kohlbergs analytische Unterscheidung zwischen Struktur und Inhalt .....	9
1.3 Fortführung des "bootstrapping" im Schnittpunkt von drei Analyseebenen .....	16
1.3.1 Programm der Reanalyse auf normativ-analytischer Ebene .....	17
1.3.2 Programm der Reanalyse auf begriffs- und bedeutungsanalytischer Ebene .....	18
1.3.3 Programm der Reanalyse auf empirisch-analytischer Ebene .....	19
2. Ergebnisse auf theoretisch-konzeptueller Ebene .....	21
2.1 Das Problem der Vermittlung von rationaler und phänomenologischer Rekonstruktion .....	21
2.1.1 Das Problem einer angemessenen Differenzierung zwischen konventionellen und postkonventionellen moralischen Urteilen .....	24
2.1.2 Das Problem angemessener Stufenbestimmungen auf konventionellem und präkonventionellem Niveau .....	33
2.1.3 Weiterentwicklung des Präferenztests von Rest .....	42
2.2 Das Problem der Vermittlung von hermeneutischer und psychometrischer Forschungsstrategie .....	43
2.3 Analyse des Standard Issue Scoring-Manuals .....	51
2.4 Das mehrdimensionale Modell .....	55
3. Ergebnisse auf empirisch-analytischer Ebene .....	63
3.1. Kohlbergs einseitige Berücksichtigung hoch aggregierter Auswertungsdaten.....	63
3.2 Ergebnisse zur Analyse der Issue-Werte .....	66
3.2.1 Der Anteil postkonventioneller Stufen bei den Issue-Werten .....	66
3.2.2 Intraindividuelle Konsistenz der Stufenwerte .....	71
3.2.3 Kontextabhängigkeit der Stufenwerte .....	76
3.3 Überprüfung der Interraterreliabilität für die Criterion Judgments .....	82
4. Zusammenfassung und Ausblick .....	93
Literaturverzeichnis .....	108

## **Zusammenfassung**

Die Evaluation des Kohlbergschen Forschungsansatzes zur Entwicklung der moralischen Urteilsfähigkeit erfolgt im Schnittpunkt von drei Analyseebenen: auf bedeutungsanalytischer, auf normativ-analytischer und auf empirisch-analytischer Ebene. Auf bedeutungs- und normativanalytischer Ebene geht es um die Klärung der Frage, ob Kohlbergs Stufenbestimmungen die Kriterien einer kohärenten Stufentheorie in kognitiv-strukturtheoretischem Sinne erfüllen, d.h. inwieweit sie hinreichend konsistent und trennscharf und in Gestalt hierarchischer Integration definiert sind. Auf empirisch-analytischer Ebene wurde überprüft, inwieweit die auf der Basis des Auswertungsmanuals (Colby et al., 1987) gewonnenen empirischen Ergebnisse mit den grundlegenden theoretischen Annahmen Kohlbergs übereinstimmen. Die Evaluation stellt einen Beitrag zu einer immanenten Kritik dar, hat darüber hinaus aber eine Weiterführung des kognitiv-strukturtheoretischen Ansatzes im Sinne produktiver Problemverschiebung (Lakatos) zum Ziel.

## **Summary**

The evaluation of Kohlberg's model of the development of moral judgment represents the interface of three different approaches: semantic, normative and empirical. The semantic and normative analyses focus on the stage theoretical aspects of the model: Do the stages fit the criteria for a coherent stage theory in a structural-developmental sense? Are they sufficiently consistent and distinctive? Do they fulfill the criteria of hierarchical integration? Empirically, the question is whether the data derived on the basis of the coding manual (Colby et al., 1987) are compatible with Kohlberg's basic theoretical assumptions. This evaluation is a contribution to an immanent critique in view of a progressive problem shift (Lakatos) serving the advancement of the cognitive-structural model.

## Evaluation des Ansatzes von Lawrence Kohlberg zur Entwicklung und Messung moralischen Urteilens

### *Vorbemerkung*

Der Versuch einer systematischen Rekonstruktion des Ansatzes von Kohlberg erfolgt im Schnittpunkt von drei Analyseebenen: auf bedeutungsanalytischer, auf normativ-analytischer und auf empirisch-analytischer Ebene. Auf bedeutungs- und normativ-analytischer Ebene geht es um die Klärung der Frage, ob Kohlbergs theoretische und operationale Stufenbestimmungen die Kriterien einer kohärenten Stufentheorie im kognitiv-strukturtheoretischen Sinne erfüllen, d.h. inwieweit sie hinreichend konsistent und trennscharf und im Sinne "hierarchischer Integration" definiert sind. Durch die Operationalisierung der Stufenbestimmungen im Kohlbergschen Auswertungsmanual wird in präjudikativer Form festgelegt, welche "möglichen Sachverhalte" empirisch überhaupt erfaßbar und nachweisbar sind. Auf empirisch-analytischer Ebene ist zu überprüfen, inwieweit die auf der Basis des Auswertungsmanuals gewonnenen empirischen Ergebnisse mit den grundlegenden theoretischen Annahmen Kohlbergs übereinstimmen.

Kohlbergs Forschungsansatz beruht auf der Wechselwirkung von bedeutungsanalytischer und normativanalytischer und empirischer Ebene. Die auf der Wechselwirkung von Empirie und präjudikativer Ebene beruhende Forschungsdynamik bezeichnet Kohlberg metaphorisch als einen Prozeß des "bootstrapping". Eine Abklärung des kognitiv-strukturtheoretischen Forschungsansatzes ist nur auf diesem Wege möglich.

Das beträchtliche Ausmaß an kohlbergkritischen Befunden, zu denen die hier vorgelegte Rekonstruktion auf den drei Analyseebenen gelangte, könnte den Eindruck erwecken, die grundlegende Untersuchungsintention sei eine vorrangig kritische. Tatsächlich hat die Rekonstruktion jedoch primär eine kritische Würdigung und Weiterführung von Kohlbergs kognitiv-strukturtheoretischem, universalistischem Forschungsansatz zum Ziel; zwischen Kohlbergs Ansatz und dessen spezifischer Durchführung und Konkretisierung in seinem Stufenmodell muß klar unterschieden werden. Kohlbergs Forschungsansatz kann nur im Rahmen eines partiell revidierten Stufenmodells in überzeugender Weise zur Geltung gebracht werden.

Kohlbergs Forschungsansatz hat allein daran seinen Maßstab, ob und inwieweit er sich als "progressives Paradigma" (Lakatos) erweist. Die Faszination dieses Ansatzes beruht dabei

darauf, daß er den Anspruch einer rationalen Rekonstruktion mit dem einer empirisch gehaltvollen, erklärungskräftigen Theorie der Entwicklung moralischer Urteilsfähigkeit verknüpft und deshalb unter beiden Theorieaspekten kritisierbar ist, aber auch seine Kraft als progressives Paradigma unter Beweis stellen kann.

Kohlbergs Forschungsansatz hat auf moralphilosophischer Ebene den Standpunkt einer autonomen und universalistischen Moral in der Tradition rationaler Normenbegründung zur Voraussetzung, ein Standpunkt, der nicht leichtfertig zugunsten modischer Trends kontextualistischer oder postmodernistischer Art preisgegeben werden sollte. Auf moralpsychologischer Ebene liegt die analoge Versuchung nahe, invariante universelle Strukturen durch kontextspezifisch variierende Strukturen zu ersetzen. Wir halten es für sinnvoll, auf moralphilosophischer wie auf moralpsychologischer Ebene an Kohlbergs starken universalistischen Annahmen festzuhalten.

Einer Würdigung des Kohlbergschen Forschungsansatzes wäre mit einer dogmatischen Festbeschreibung der von Kohlberg vorgenommenen Stufenbestimmungen - der rein theoretischen wie der operationalen im Auswertungsmanual - am schlechtesten gedient. Kohlbergs Forschungsansatz kann sich als progressives Paradigma nur dann erweisen, wenn die Durchführung, die er in Kohlbergs Stufenmodell erfahren hat, an seinen eigenen, konstitutiven Ansprüchen gemessen wird. Es war deshalb zu untersuchen, inwieweit der Anspruch einer Vermittlung von rationaler und phänomenologischer Rekonstruktion, aber auch von hermeneutischem und psychometrischem Forschungsansatz bei Kohlberg erfüllt worden ist. Außerdem war die Frage zu klären, inwieweit Kohlbergs Stufenmodell tatsächlich auf einer klaren und angemessenen Trennung zwischen den Struktur- und Inhaltsaspekten moralischen Urteilens beruht.

Eine Würdigung des Kohlbergschen Forschungsansatzes als eines progressiven Paradigmas macht methodologisch zwei Schritte erforderlich: einen immanent kritischen und einen konstruktiven. Im Rahmen der Rekonstruktion waren deshalb einerseits auf den o.g. drei Analyseebenen die Problemzonen kenntlich zu machen, die für Kohlbergs Stufenbestimmungen charakteristisch sind. Zum anderen aber mußte der ungleich schwierigere Versuch unternommen werden, Gesichtspunkte (Konzepte) zu entwickeln sowie Wege (Methoden) aufzuzeigen, die es erlauben, diese Problemzonen in konstruktiver Weise zu bewältigen.

Die hier aufgewiesenen Wege fallen in das Paradigma der Wechselwirkung von internen Zwängen und externen Bedingungen, die die Entwicklungsdynamik der kognitiven wie der moralischen Entwicklung zu rekonstruieren erlauben (Edelstein, 1992; Edelstein & Hoppe-Graff, 1993).

Das Projekt Kohlbergrekonstruktion, dessen Ergebnisse hier berichtet werden, wurde von der DFG in Kooperation zwischen einem Philosophen und einem Psychologen - also unter philosophischer und psychologischer Perspektive - gefördert. Innerhalb dieses Projekts wurden Kohlbergs Daten im Rahmen des interdisziplinären Forschungsprogramms "Philosophische Ethik - Interdisziplinärer Ethikkurs" reanalysiert und die Ergebnisse mit den Befunden unserer eigenen Längsschnittdaten im Projekt "Entwicklung und Soziale Struktur" sowie mit den Daten unserer kulturvergleichenden Studien in der VR China, Rußland sowie in Ost- und Westdeutschland verglichen. Das so entstandene Archiv moralpsychologischer Längsschnittdaten enthält wichtige Anregungen für eine moralpsychologische Forschung auf breiter Grundlage. Der deutschen Forschungsgemeinschaft sei für die Ermöglichung dieser Forschung an dieser Stelle gedankt.

Wolfgang Edelstein

## 1. Kohlbergrekonstruktion im Schnittpunkt von drei Analyseebenen

Gegenstand unserer Reanalyse ist die von Kohlberg in den USA durchgeführte Längsschnittuntersuchung zur Ontogenese der moralischen Urteilsfähigkeit. Die Datenerhebung erfolgte in Form struktureller Interviews, und zwar über 6 Meßzeitpunkte im Abstand von 3 bis 4 Jahren. Kohlbergs Längsschnittstudie nahm in den USA Ende der 50er Jahre ihren Ausgang und erstreckte sich über mehr als 20 Jahre. Sie geht von drei Kohorten 10-, 13- und 16jähriger Kinder bzw. Jugendlicher aus und umfaßt eine Stichprobe von 58 Personen, freilich nur männlichen Geschlechts, die sowohl der oberen und unteren Mittelschicht als auch der Unterschicht angehören. Die Beschränkung auf männliche Personen ist zweifellos ein gravierender Mangel, da man ja kein Wissen a priori darüber besitzt, ob bzw. inwieweit geschlechtsspezifische Unterschiede in den Lebens- und Arbeitsverhältnissen sich entweder als Konstitutions- und Performanzbedingungen auf die Entwicklung von invarianten, universellen Tiefenstrukturen der moralischen Urteilsfähigkeit auswirken (s. Edelstein, 1992) oder aber zu alternativen Entwicklungspfaden führen (s. Gilligan, 1984). Trotz dieses Defizits stellt Kohlbergs Längsschnittuntersuchung eine der umfassendsten im Bereich der Entwicklungspsychologie überhaupt und die bedeutendste im moralpsychologischen Bereich dar. Die Längsschnittuntersuchung liefert die empirische Basis für Kohlbergs Stufenmodell, das - ungeachtet der Revisionen, die es durch Kohlberg selber in einer 30 Jahre umspannenden Forschungsgeschichte erfuhr - innerhalb des kognitiv-strukturtheoretischen "Lagers" maßgeblich blieb.

### 1.1 Kohlbergs Interview- und Auswertungsverfahren

Kohlbergs Daten sind Aussagen zu Gerechtigkeitskonflikten, bei denen nicht Recht mit Unrecht kollidiert - so wäre die Entscheidung eine von vornherein klar vorgezeichnete - sondern Recht auf Recht trifft. (Im Heinz-Dilemma steht z.B. dem Recht auf Leben, das Heinz in Identifizierung mit den Interessen und Ansprüchen seiner krebskranken Frau geltend zu machen sucht, das Eigentumsrecht eines Apothekers gegenüber). Die Aussagen wurden längsschnittlich, von Kindheit und Jugend bis ins reife Erwachsenenalter, anhand eines halbstandardisierten strukturellen Interviews (moral judgment interview) gewonnen und mittels eines standardisierten Auswertungsverfahrens (standard issue scoring) ausgewertet.

Beide Instrumente haben im Laufe der Entwicklung von Kohlbergs Forschung eine Veränderung erfahren.



In das moralische Interview wurden zunehmend konzeptuelle Fragen ("warum ist es wichtig, dem Gesetz zu gehorchen", "ein Versprechen zu halten" usf.) aufgenommen. Außerdem wurde die Interviewführung zu den späteren Meßzeitpunkten strenger an die standardisierten Fragen gebunden; zugleich wurden konsequenter Nachfragen im Sinne von Warum-Fragen und Erläuterungsfragen gestellt. - Kohlberg ging es bei der Datenerhebung nicht um die Erfassung von moralischen Einstellungen oder Überzeugungen, sondern um die Rekonstruktion der invarianten Strukturen moralischer Urteilsfähigkeit, die auf der Differenzierung zwischen mehr oder weniger tragfähigen Begründungsformen sowohl (a) für moralische Entscheidungen als auch (b) für die Verbindlichkeit moralischer Normen beruht. Punkt (a) ist dabei auf situationsspezifische Fragen (z.B. "soll Heinz das für seine Frau lebenswichtige Medikament stehlen und warum?"), Punkt (b) hingegen ist auf konzeptuelle Fragen bezogen (z.B. "warum ist es wichtig, dem Gesetz zu gehorchen"). Die Interviews, in denen diese Gesichtspunkte zur Geltung gebracht werden, beziehen sich auf hypothetische Dilemmata. Diese hat Kohlberg realistischen Dilemmata (real life dilemmas) vorgezogen, weil sie eine größere Vergleichbarkeit hinsichtlich der konstitutiven Interessen- und Normkonflikte garantieren und weil er annahm, daß sie das zu erwartende Kompetenz-Performanzgefälle minimieren. Im übrigen repräsentieren sie das testtheoretische Prinzip gleicher Bedingungen (equally unfair) für die Messung.

Auf der Ebene des Auswertungsinstruments wurden drei Varianten entwickelt, nämlich das "Sentence Rating and Global Story Rating" (Ende der 50er Jahre), das "Structural Issue Scoring" (Anfang der 70er Jahre) und das "Standard Issue Scoring" (Ende der 70er Jahre). Kohlberg schreibt nur dem Standard Issue Scoring eine hinreichend hohe Validität und Reliabilität zu. Das Sentence Rating and Global Story Rating hingegen, bei dem noch nicht analytisch zwischen den Struktur- und Inhaltsaspekten moralischen Urteilens unterschieden wurde, erscheint ihm weder valide noch reliabel, während das Structural Issue Scoring zwar valide, aber nur beim Einsatz gut geschulter und mit der Theorie vertrauter Rater hinreichend reliabel ist: "In Kohlberg's early work ... stages were assessed using an interview based on hypothetical moral dilemmas and two systems for scoring responses to those interviews: Sentence Rating and Global Story Rating. Some 12 years later, a new scoring system, Structural Issue Scoring, was introduced. From a psychometric point of view, there were serious problems with these scoring systems. The early systems lacked validity as measures of structural-developmental stages since they involved what was essentially a content analysis of interview responses. The later system achieved increased validity when used by very experienced raters but was subjective, difficult to use, and unreliable, particularly when used by those who had not had extensive training and who had not studied with Kohlberg. Data will be presented here to show that the current system, Standard Issue Scoring, provides for

reliable and valid assessment of moral judgment stage. This was achieved through redefining the unit of analysis and the relation between moral judgment content and structure and through specifying more precisely the process of inference from interview material to stage scores" (Colby et al., 1983, S. 13).

In Kohlbergs Interviewverfahren (Moral Judgment Interview) werden 9 Standarddilemmata berücksichtigt, wobei jeweils 3 zu einer Form zusammengefaßt sind (Form A umfaßt das Heinz- und das Joe-Dilemma sowie "Officer Brown", Form B das Euthanasie- und das Judy-Dilemma sowie "Dr. Rogers", Form C das Korea- und Jean-Valjean-Dilemma sowie das Bob-Karl-Dilemma). Jeweils drei Dilemmata repräsentieren dabei 3 ähnliche Paare konfligierender Wertorientierungen, die sog. Issues (nämlich "Leben" vs. "Gesetz", "Gewissen" vs. "Strafe" und "Versprechen" vs. "Autorität"). Die gegen eine Autorität oder gegen das Gesetz gerichtete Entscheidung (action choice) wird als autonome, die entgegengesetzte Option hingegen als heteronome verstanden.

Obwohl die Zusammenfassung bestimmter Dilemmata zu einer Form grundsätzlich beliebig ist, sofern nur die Voraussetzung erfüllt ist, daß in jeder Form die drei Issue-Paare vertreten sind, erfolgte die Zusammenstellung de facto doch in der Weise, daß die Formen A und B gegenüber C in stärkerem Maße postkonventionelle Werte aufweisen, die Formen A und B unter dem Gesichtspunkt der Manifestation bzw. der Erfäßbarkeit der höchsten Kompetenz gegenüber C also den Vorzug verdienen.

Den zentralen, maßgeblichen Bestandteil des Standard Issue Scoring stellt das "Standard Issue Scoring-Manual" dar, das ein atomistisches Auswertungsverfahren repräsentiert: isolierte, einzelne Intervieweinheiten (interview judgments) werden prototypischen Äußerungen - sog. "Criterion Judgments" - als isolierten, dekontextualisierten Auswertungseinheiten zugeordnet. Eine Darstellung des komplizierten Filtersystems, das passiert werden muß, bzw. der 17 Schritte, die diese Zuordnung regeln, würde im Rahmen dieser Einführung zu weit führen (vgl. Colby & Kohlberg, 1987, S. 158-188).

Kohlbergs Auswertungsmanual beruht nicht auf rein konzeptuellen Überlegungen a priori, sondern hat auch eine empirische Grundlage. Es wurde nämlich auf der empirischen Basis von 7 nach dem Zufallsprinzip ausgewählten Fällen konstruiert, die zu allen 6 Meßzeitpunkten interviewt wurden. Die 7 Fälle wurden zunächst einmal anhand des Structural Issue Scoring in diskursiver Form ausgewertet und erhielten einen globalen Stufenwert. Auf dieser Basis wurde eine "konzeptuelle Analyse" vorgenommen, die eine differenziertere Auswertung im atomistischen Sinne, d.h. hinsichtlich der durch isolierte,

einzelne Criterion Judgments (CJs) repräsentierten Stufenwerte ermöglichen sollte: "Seven cases were selected from Kohlberg's longitudinal sample. These 'construction cases', chosen at random from among those tested at all six times, were assigned global stage scores based on intensive discussion and analysis using concepts from Structural Issue Scoring. The responses to each dilemma were then classified into clearly defined scoring units, or "Interview Judgments". Each of these interview judgments formed the basis for a criterion judgment to be entered in the scoring manual. The stage score of each criterion judgment was assigned on the basis of the global score of the interview from which it was derived and a conceptual analysis of the idea it embodied. The criterion judgments generated by these seven construction cases were later used to score the remaining interviews in the longitudinal study through a process of matching interview material to criterion judgments in the manual. Those cases not in the construction sample comprised a blind sample which was not used at all until the scoring manual had been completed" (Colby et al., 1983, S. 8f.).

Die im Prozeß der Auswertung ermittelten, durch CJs repräsentierten Stufenwerte stehen - trotz des atomistischen Auswertungsverfahrens - der Sichtweise der Subjekte selbst noch relativ nahe. Um zu globalen Aussagen über das Entwicklungsniveau bei bestimmten Issues zu kommen, werden CJ-Werte nach bestimmten Verrechnungsregeln zu "Issue-Werten" zusammengefaßt. Dilemma-Werte werden nicht gebildet, vielmehr werden die Issue-Werte zu "Form-Werten" zusammengefaßt.

## 1.2 Kohlbergs analytische Unterscheidung zwischen Struktur und Inhalt

Bis Ende der 60er Jahre war Kohlbergs Stufenmodell nicht auf einer klaren analytischen Distinktion zwischen den "Strukturen" und "Inhalten" moralischen Urteilens aufgebaut; vielmehr wurden zur Stufenkennzeichnung bestimmte Inhaltsaspekte verwendet, genauer und schärfer gesagt: die Stufen wurden mit bestimmten Inhalten identifiziert (z. B. ein law-and-order-Standpunkt mit Stufe 4 und Standards eines geschärften individuellen Gewissens mit Stufe 5).

Die Konfundierung von Struktur- und Inhaltsaspekten moralischer Urteile führte jedoch - wie die Studie von Kramer (1968) nachwies - zu (scheinbaren) Stufenregressionen in einem Umfang, der mit Kohlbergs kognitivistischem Stufenmodell nicht mehr zu vereinbaren war. Kohlbergs kognitivistische Position steht oder fällt ja mit der grundlegenden Behauptung, daß hinter die einmal erreichte Problemlösungskapazität bzw. Einsicht nicht mehr zurückgegangen werden kann, jedenfalls nicht aus kognitiv-internen, d.h. rationalen Gründen.

In Reaktion auf die Studie von Kramer versuchte Kohlberg in den 70er Jahren, ein revidiertes Stufenmodell zu entwickeln, das auf der Ebene theoretischer wie operationaler Stufenbestimmungen auf einer klaren analytischen Unterscheidung zwischen den Strukturen und Inhalten moralischer Urteilsbildung beruht.

Das Konzept der "soziomoralischen Perspektive" stellt - auf der Ebene der theoretischen wie der operationalen Stufenbestimmungen - das grundlegende, stufenkonstitutive Strukturmerkmal dar. Dieses Konzept ist als ein genuin moralpsychologisches zu verstehen und nicht im Sinne einer bloßen Anwendung von Formen der Perspektivendifferenzierung- und -koordination, wie sie bereits auf der Ebene deskriptiver sozialer Kognition (vgl. Keller, 1976; Selman, 1984) erworben werden bzw. rekonstruierbar sind, auf den Bereich moralischer Urteilsfähigkeit.

Die Entwicklung soziomoralischer Perspektiven wird i. S. "hierarchischer Integration" interpretiert: das Verständnis und die Lösung moralischer Konflikte erfolgt z.B. auf Stufe 3 unter einer (interpersonellen) "Beziehungsperspektive", die auf Stufe 4 im Lichte einer komplexeren "Systemperspektive" selber zum Gegenstand der Reflexion gemacht werden kann.

Im Lichte einer soziomoralischen Perspektive können die begründungsrelevanten Gegebenheiten einer Konfliktsituation und die damit verknüpfbaren moralischen Bedeutungen stufenspezifisch in einem einheitsstiftenden Sinne organisiert und aufeinander bezogen werden, und zwar in Form von Perspektivenübernahmen, die im Zuge einer invarianten Entwicklungssequenz moralischen Urteilens zunehmend komplexer, äquilibrierter und reversibler werden. Die durch Stufe 6 als ausgezeichnetem normativem Bezugspunkt repräsentierte soziomoralische Perspektive kennzeichnet Kohlberg als "ideal role-taking", was soviel wie eine vollständig reversibel und universalisierbar gewordene Perspektivenübernahme bedeutet.

Die Reversibilität und Äquilibration moralischen Urteilens ist an der (mehr oder weniger umfassenden) intersubjektiven Verbindlichkeit bzw. Universalisierbarkeit von moralischen Gesichtspunkten festgemacht, die alle potentiell rationalen Subjekte miteinander teilen und in deren Lichte überhaupt erst ein angemessenes Verständnis und eine angemessene Berücksichtigung bzw. die erforderliche Transformation der konfligierenden partikularen Interessenstandpunkte, Norm- und Wertorientierungen erfolgen kann.

Die zunehmend komplexer, äquilibrierter und reversibler werdenden Formen der Perspektivenübernahme ermöglichen ein moralisches Konfliktverständnis und Konfliktlösungen, die in wachsendem Maße die Zustimmung aller Betroffenen finden (z.B. auf Stufe 2 im Sinne einer ausbalancierten Berücksichtigung von individuellen Eigeninteressen und auf Stufe 3 in Form eines hypothetischen Rollentauschs, wie er mit der "konkreten Goldenen Regel" impliziert ist).

Die Übernahme soziomoralischer Perspektiven durchläuft 6 Stufen, wobei jeweils zwei Stufen zu einem Entwicklungsniveau zusammengefaßt sind; "Niveaus" definieren sich durch den Standpunkt, den moralische Subjekte gegenüber der sozialen Gemeinschaft und ihren etablierten Standards einnehmen.

(1) Auf präkonventionellem Niveau (Stufen 1 und 2) wird der Standpunkt eines "isolierten Individuums" bezogen, das auf Stufe 1 noch in einer egozentrischen Perspektive befangen bleibt, d.h. keine Differenzierung zwischen unterschiedlichen Interessen und Ansprüchen von Ego und Alter vornimmt. Auf Stufe 2 hingegen kommt es zu einer Differenzierung und Koordination unterschiedlicher Interessenlagen; deren Berücksichtigung erfolgt jedoch im instrumentalistisch-egoistischen Sinne in Form einer Relativierung des "Richtigen" auf unterschiedliche Eigeninteressen.

(2) Auf konventionellem Niveau (Stufen 3 und 4) wird zwar der Standpunkt eines isolierten Individuums durch den Standpunkt von Individuen als Mitgliedern einer sozialen Gemeinschaft überwunden, die damit implizierte, vorrangige Berücksichtigung gemeinsamer Interessen, an denen die erforderliche Transformation bzw. Rechtfertigbarkeit von Eigeninteressen ihren moralischen Maßstab hat, bleibt jedoch an die Grenzen partikularer Gruppenmoral gebunden. Die etablierten Standards partikularer Gruppenmoral können noch nicht von einem übergeordneten, universalisierbaren Standpunkt der Moral aus hinterfragt und transzendiert werden. Dies gilt für Stufe 3 und Stufe 4 gleichermaßen. Für Stufe 3 ist eine "Beziehungsperspektive" konstitutiv, d.i. die Perspektivenübernahme eines Individuums, das konfligierende Rechte und Pflichten im Rahmen konkreter interpersoneller Beziehungen versteht, in denen es mit anderen Individuen steht. Die Rechtfertigung moralischer Handlungsentscheidungen und Normen geschieht im Wissen um berechnete Erwartungen und Motive, die Individuen im Rahmen partikularer Gruppenmoral miteinander teilen. Die Entwicklung des für Stufe 3 kennzeichnenden Normverständnisses bleibt dabei an konkrete, interpersonelle Beziehungen gebunden. Es bleibt folglich ein beziehungspezifisches Normverständnis. Auf Stufe 4 wird eine "Systemperspektive" eingenommen, die es erlaubt, unterschiedliche interpersonelle Beziehungen und die in ihnen geltenden Regeln und

Rollenerwartungen im Hinblick auf soziale Institutionen und die generellen Funktionsbedingungen sozialer Systeme zu verstehen und zu berücksichtigen.

(3) Erst auf postkonventioneller Ebene (Stufen 5 und 6) erfolgt eine Differenzierung und Koordination von unterschiedlichen Interessen und Ansprüchen im Lichte eines universalisierbaren Standpunkts der Moral. Dieser Standpunkt trägt vorrangig dem Rechnung, was gleichermaßen im Interesse aller Individuen als bedürftiger und verletzbarer vernunftgeleiteter Menschen liegt. Diese versuchen, ihre versehrbare Integrität als moralische Subjekte, d.h. als Zentren eigener, selbstbestimmter Urteils- und Willensbildung dadurch zu schützen, daß sie im Rekurs auf universelle Grundrechte in gegenseitiger Achtung und Toleranz um Formen der Konfliktlösung bemüht sind, die in gleicher Weise im Interesse aller Betroffenen liegen. Die unter dem "moralischen Gesichtspunkt" gebotene Transformation partikularer Interessen und Standards (individuell selbstgesetzter wie sozial etablierter Art) hat deshalb ihren Maßstab an Verfahren "der Normierung von Normen", d.h. der Normüberprüfung, Normenkonstitutierung und -durchsetzung, die als legitime Verfahren nur dann zu verstehen sind, wenn sie die Zustimmung aller Betroffenen finden können. Für Stufe 5 ist die soziomoralische Perspektive eines rationalen Individuums konstitutiv, das an Werten und Grundrechten orientiert ist, die sozialen Bindungen und Vereinbarungen, soweit sie an die etablierten Standards partikularer Gruppenmoral gebunden bleiben, übergeordnet sind. Soziale Systeme werden danach bewertet, inwieweit sie die Menschenrechte und das Wohl aller Individuen schützen. Für Stufe 6 ist eine Perspektivenübernahme kennzeichnend, die vollständig reversibel und universalisierbar geworden ist. Nur unter dieser Voraussetzung ist nämlich eine denkbar höchste Entwicklungsstufe moralischer Urteilsfähigkeit antizipierbar, die mit dem Standpunkt einer autonomen und universalistischen Moral in der Tradition rationaler Normenbegründung in Übereinstimmung steht. Die Perspektivenkoordination erfolgt durch formale Mechanismen bzw. Verfahrensregeln.

Kohlbergs Stufenbestimmungen sind in didaktisch überzeugender Form in dem von Oser & Althof verfaßten Lehrbuch "Moralische Selbstbestimmung" (1992) dargestellt worden.

Das methodologische Postulat, die (analytische) Unabhängigkeit der Strukturen von den Inhalten hinreichend zu berücksichtigen, ist auf der operationalen Ebene der Stufenbestimmungen anders zu verstehen als auf der theoretischen. Auf theoretischer Ebene wird versucht, die Stufen durch reine Strukturmerkmale zu kennzeichnen. Die durch die "criterion judgments" (CJs) im Auswertungsmanual repräsentierten Stufenbestimmungen sind hingegen per definitionem durch die Verknüpfung von Strukturaspekten mit drei

Inhaltsdimensionen gekennzeichnet ("Entscheidungsrichtung" sowie generelle "Normen" und spezifische "Elemente"), auf denen Entscheidungsbegründungen inhaltlich beruhen.

Kohlberg operiert mit 9 "Normen" (z.B. Wahrung des Lebens, Verbindlichkeit von Gesetzen, eines Versprechens, des Rechts von Autoritätspersonen), von denen er annimmt, daß sie in allen Kulturen auftreten. Das gleiche gilt für die "Elemente"-Gruppen, zwischen denen er differenziert. Kohlberg führt 17 Elemente auf, die er in 4 Gruppen zusammenfaßt:

1. Gruppe: Aufrechterhaltung der normativen Ordnung (Elemente: 1."Gehorsam", 2."Tadeln", 3."Vergelten", 4."ein Recht haben", 5."eine Pflicht haben").
2. Gruppe: Berücksichtigung von Handlungsfolgen (5a."egoistisch-utilitaristische Orientierung", 5b."sozialutilitaristische Orientierung": "positive Konsequenzen für das Individuum", "positive Konsequenzen für die Gruppe", 6."gutes Ansehen", 7."Belohnung").
3. Gruppe: perfektionistische Orientierung an Idealen des Selbst und des sozialen Zusammenlebens (10."Aufrechterhaltung des Charakters", 11."Aufrechterhaltung des Respekts vor sich selbst", 12."dient einem sozialen Ideal oder sozialer Harmonie", 13."dient der menschlichen Würde und Autonomie").
4. Gruppe: Orientierung an Fairneß-Gesichtspunkten (14. "Ausbalancierung von Perspektiven", 15. "Reziprozität oder positive Vergeltung", 16. "Aufrechterhaltung von Unparteilichkeit", 17. "Aufrechterhaltung sozialer Kontrakte").

Auf theoretischer Ebene werden generelle, d.h. dilemma-, issue- und fragenübergreifende Stufenbestimmungen vorgenommen. Von einer systematischen Verknüpfung stufenspezifischer Strukturmerkmale und Inhaltsaspekte, wie sie sich in den drei Inhaltsdimensionen definieren, wird abstrahiert. Auf der Ebene operationaler Stufenbestimmungen hingegen definieren sich die für bestimmte Stufen konstitutiven Strukturmerkmale in den drei Inhaltsdimensionen und sind dadurch gekennzeichnet, daß sie bestimmte Inhalte (Elemente) ein- bzw. ausschließen.

Auf präkonventionellem Niveau schließt z.B. die für Stufe 1 kennzeichnende "heteronome Moral" die erste und zweite Elementengruppe ("Aufrechterhaltung der normativen Ordnung" und "Berücksichtigung von Handlungsfolgen") ein, wohingegen die dritte und vierte Elementengruppe ("perfektionistische Orientierung an Idealen des Selbst und des sozialen Zusammenlebens" sowie die "Orientierung an Fairneß-Gesichtspunkten") weitgehend ausgeschlossen ist.

Kohlberg hat versucht, im Rahmen des auf der Distinktion von Struktur und Inhalt beruhenden Stufenmodells eine typologische Differenzierung vorzunehmen. Anfang der 70er Jahre wurde von ihm zunächst ein Unterstufen-Konzept entwickelt. Dabei wurde zwischen

einer autonomen Unterstufe (B) und einer heteronomen Unterstufe (A) unterschieden. Die Unterstufe A wird durch die beiden ersten Elementgruppen definiert ("Aufrechterhaltung der normativen Ordnung" und "Berücksichtigung von Handlungsfolgen"), während die Unterstufe B durch die beiden anderen Elementgruppen definiert ist ("perfektionistische Orientierung an Idealen des Selbst und des sozialen Zusammenlebens" sowie "Orientierung an Fairneß-Gesichtspunkten").

Anfang der 80er Jahre wurde das Unterstufen-Konzept aufgegeben. An seine Stelle trat ein Konzept idealer Typen. Die typologische Unterscheidung zwischen autonomer und heteronomer Orientierung wurde zwar beibehalten, erfolgte aber auf der Grundlage anderer Kriterien, nämlich von folgenden neun Typenkriterien: 1. Entscheidungsrichtung, 2. Autonomie, 3. gegenseitiger Respekt, 4. Reversibilität, 5. Konstruktivismus, 6. Wertehierarchisierung, 7. intrinsische Wertorientierung, 8. Präskriptivität, 9. Universalisierung. - Kohlberg versuchte, diese neun Kriterien für jedes Dilemma eigenständig zu operationalisieren. Die bei einem Dilemma hinsichtlich dieser neun Kriterien erzielten Werte werden zu einem Dilemmawert verrechnet und die drei Dilemmawerte einer Form zu einem Formwert zusammengefaßt (vgl. Colby & Kohlberg, 1987, S. 315ff.)

Das Unterstufen-Konzept hat gegenüber dem späteren Konzept idealer Typen den formalen Vorzug, daß es klar zwischen Struktur- und Inhaltsaspekten zu trennen sucht. Regressionen (von B zu A) innerhalb einer Stufe werden ausgeschlossen, während ein Überspringen von Unterstufe B eingeräumt wird. Das Unterstufenkonzept konnte freilich empirisch nicht bestätigt werden. Das spätere Typenkonzept hingegen, das von einer "tendenziellen Entwicklung" vom heteronomen zum autonomen Typus ausgeht (s. Colby & Kohlberg, 1987, ebd.), konnte weitgehend bestätigt werden.

Die starke Hypothese des ersten Typenkonzepts beruhte konzeptuell auf der von vornherein problematischen, weil rein mechanischen Definition der Unterstufe A durch die beiden ersten Elementengruppen und der Unterstufe B durch die beiden anderen Elementgruppen.

Das spätere Typenkonzept weist konzeptuell ebenfalls Mängel auf. Unter bedeutungsanalytischen Gesichtspunkten ist es nämlich fragwürdig, ob bzw. inwiefern es sich (a) bei den (positiv bestimmten) Kriterien des Typus B um Kriterien handelt, die voneinander unabhängig sind sowie (b) um Kriterien, die mit den Standards postkonventionellen Urteilens nicht konfundiert sind (die Kriterien "intrinsische Wertorientierung", "Wertehierarchisierung" und "Universalisierung" erweisen sich unter den Gesichtspunkten (a)



und (b) als problematisch). Außerdem ist Typus A unterbestimmt; er definiert sich nur per negationem vom Typus B her.

Sowohl die theoretischen als auch die operationalen Stufenbestimmungen können nur dann als angemessen betrachtet werden, wenn sie folgende fünf formale Kriterien erfüllen, an denen Kohlberg selbst orientiert ist:

(1) Es darf keine Konfundierung der Struktur- und Inhaltsaspekte bzw. keine inhaltliche Verengung der Stufenbestimmung auftreten. (Dies ist z.B. bei Stufe 2 der Fall, wenn diese im Sinne einer egoistisch-instrumentalistischen Sichtweise verstanden wird (vgl. Keller, Eckensberger & v. Rosen, 1989; Keller, 1990). Die Vermeidung einer Konfundierung bzw. inhaltlichen Verengung ist zwar - als methodologisches Postulat - mit der grundbegrifflichen, analytischen Trennung zwischen den Strukturen und Inhalten moralischer Urteile impliziert, ist aber keine auf rein begriffsanalytischer Ebene entscheidbare Frage. Inwiefern Konfundierungen bzw. inhaltliche Verengungen auftreten, kann nur im Schnittpunkt von konzeptueller und empirischer Analyse überprüft werden.

(2) Die Stufenbestimmungen müssen sich als hinreichend konsistent und trennscharf erweisen.

(3) Die Stufenbestimmungen müssen das Kriterium "hierarchischer Integration" erfüllen. Dieses Kriterium interpretiert die voraussetzungslogische Beziehung zwischen der soziomoralischen Perspektive und den Standards einer Stufe (n) in Relation zu Stufe (n+1) in der Weise, daß (n) in eine komplexere Perspektive (n+1) integriert wird, die im Vergleich zu (n) eine äquilibriertere und reversiblere Form der Perspektivenübernahme ermöglicht.

(4) Die Stufenbestimmungen müssen eine invariante, universelle Stufensequenz repräsentieren. Dieses Kriterium ist zwar mit unterschiedlichen Verlaufstypen moralischer Urteilsbildung kompatibel, sofern diese auf einer Fokussierung auf bestimmte Inhalte beruhen, jedoch nicht vereinbar mit der Annahme und dem Nachweis alternativer Entwicklungspfade (z.B. bereichsspezifischer, kultur- oder geschlechtsspezifischer Art). Kohlberg versteht das Kriterium der "invarianten Sequenz" in dem starken voraussetzungslogischen Sinne, daß damit das "Überspringen" von Stufen theoretisch, d.h. aus kognitiv-internen, rationalen Gründen ausgeschlossen ist.

5. Das Kriterium der "strukturierten Ganzheit" bzw. der internen Konsistenz, mit der im Zuge der Entwicklung von einer Stufe (n) Gebrauch gemacht wird, ist das schwächste Kriterium. Vom (relativ) konsistenten Gebrauch einer Stufe (n) ist bei Kohlberg nämlich auch dann die Rede, wenn nicht nur von (n), sondern von (n) in Verknüpfung mit einer unmittelbar benachbarten (Übergangs)-Stufe (n-1) oder (n+1) Gebrauch gemacht wird (treten drei oder mehr Stufen auf, so wird das Auftreten dieser Stufen als Meßfehler interpretiert, sofern deren Anteil mehr als 10% nicht übersteigt).

### 1.3 Fortführung des "bootstrapping" im Schnittpunkt von drei Analyseebenen

Kohlbergs grundbegriffliche Unterscheidung zwischen Struktur und Inhalt sowie die in diesem grundbegrifflichen Rahmen vorgenommenen theoretischen und operationalen Stufenbestimmungen stellen das Resultat eines "bootstrapping"-Prozesses dar. Das Verfahren des bootstrapping versteht sperrige, nicht-theoriekonforme Befunde nicht schlicht als Falsifikationsinstanz, sondern überprüft unter dem Gesichtspunkt der Angemessenheit (a) theoretischer und (b) operationaler Stufenbestimmungen, inwiefern die nicht-theoriekonformen empirischen Befunde primär in grundbegrifflich-analytischen Defiziten des Theorierahmens, etwa in der mangelnden Distinktion zwischen den Struktur- und Inhaltsaspekten moralischer Urteile oder auch in Defiziten bei der Operationalisierung theoretischer Stufenbestimmungen, d.h. in einer mangelhaften Standardisierung des Auswertungsmanuals begründet sind.

Kohlberg hält den Forschungsprozeß des bootstrapping für abgeschlossen. Er geht von der Gültigkeit bzw. Angemessenheit der von ihm in den frühen 80er Jahren selbst noch der wissenschaftlichen Öffentlichkeit vorgestellten theoretischen und operationalen Stufenbestimmungen aus.

Eine Weiterentwicklung der von Kohlberg geleisteten Forschungsarbeit steht vor zwei alternativen Möglichkeiten: (a) sie verhält sich zu den von Kohlberg auf theoretisch-konzeptueller Ebene vorgenommenen Stufenrevisionen und deren Operationalisierung im Auswertungsmanual affirmativ, hat diese also zur nicht mehr hinterfragten Voraussetzung. In diesem Fall erfolgt die Weiterentwicklung auf rein empirisch-analytischer Ebene und unter vorwiegend quantitativem (psychometrischem) Aspekt (z.B. bei Krebs et al., 1991 und Walker, 1989). Oder aber (b), die Weiterentwicklung überprüft vorgängig die von Kohlberg vorgenommenen theoretisch-konzeptuellen Revisionen sowie deren Operationalisierung,

betrachtet also das bootstrapping als einen noch nicht abgeschlossenen Prozeß (Keller, 1990; Keller, Eckensberger & von Rosen 1989).

Wir gelangen im Rahmen unserer Analysen, die Kohlbergs Stufenmodell an seinen eigenen Ansprüchen messen und insofern immanent kritisch bleiben, zu dem Ergebnis, daß sowohl die theoretischen als auch die operationalen Stufenbestimmungen Kohlbergs teilweise problematisch und revisionsbedürftig sind.

Vorgängig hatte die Analyse die sich auf bedeutungs- und normativ-analytischer Ebene stellenden Fragen zu klären, ob bzw. inwiefern Kohlbergs Stufenmodell die Kriterien einer kohärenten Stufentheorie erfüllt. Diese Kriterien fordern (a) eine klare Trennung zwischen den Struktur- und Inhaltsaspekten moralischer Urteile; (b) hinreichend konsistente und trennscharfe Stufenbestimmungen und (c) Standards rationaler Rekonstruktion, die es erlauben, angemessen zwischen mehr oder weniger intersubjektiv verbindlichen bzw. universalisierbaren Begründungsformen für moralische Entscheidungen und Normen zu unterscheiden.

### *1.3.1 Programm der Reanalyse auf normativ-analytischer Ebene*

Auf normativ-analytischer Ebene wird versucht, den Nachweis zu erbringen, daß Kohlbergs Stufenbestimmungen das Kriterium "hierarchischer Integration" nicht hinreichend erfüllen und auf der Ebene der theoretischen wie auch der operationalen Stufenbestimmungen nicht angemessen zwischen unterschiedlichen postkonventionellen Begründungsformen differenzieren, die universalistisch sind, aber autonomer oder nicht-autonomer Art sein können, bzw. autonom sind, aber universalistischer oder nicht-universalistischer Art sein können. Diese konzeptuellen Differenzierungen werden im Abschnitt "Das Problem der Vermittlung von rationaler und phänomenologischer Rekonstruktion" konkret entfaltet und begründet (S. 21ff.).

Auf dieser Analyseebene wurde auch das normative Konzept "situationsadäquater Urteilsbildung" entwickelt. Dieses Konzept erlaubt es, auch eine beträchtliche (z.B. von Stufe 3 bis Stufe 5 reichende) Inkonsistenz der Argumentationsstufen, die sinnvollerweise nicht mehr im Sinne des von Piaget entlehnten Kriteriums "strukturierter Ganzheit" interpretierbar ist, als eine kohärente Sinn Ganzheit zu verstehen. Das Konzept situationsadäquater Urteilsbildung wird in dem Abschnitt "Das Problem der Vermittlung von hermeneutischer und psychometrischer Forschungsstrategie" näher erläutert (S. 42ff.).

Primär unter normativ-analytischen Gesichtspunkten wurden auch die Dimensionen abgeklärt, die für das von uns unten eingeführte "mehrdimensionale Modell" konstitutiv sind. In diesem - handlungstheoretisch angelegten - Stufenmodell wird zwischen folgenden Dimensionen moralischer Urteils- bzw. Handlungsfähigkeit differenziert:

- (a) Formen der Handlungskoordination
- (b) Formen und Inhalten präskriptiver sozialer Kognition vs. Formen und Inhalten deskriptiver sozialer Kognition
- (c) Geltungs- vs. Anwendungsfragen
- (d) deontische Urteile, Verantwortlichkeitsurteile

Diese Dimensionen werden in dem Abschnitt "das mehrdimensionale Modell" dargestellt (S. 53ff).

### *1.3.2 Programm der Reanalyse auf begriffs- und bedeutungsanalytischer Ebene*

Auf begriffs- bzw. bedeutungsanalytischer Ebene wurde von uns untersucht, inwieweit Kohlbergs Stufenbestimmungen auf einer klaren analytischen Trennung zwischen den Struktur- und Inhaltsaspekten moralischen Urteilens beruhen. Sind die Strukturmerkmale und die für sie konstitutiven soziomoralischen Perspektiven angemessen, d.h. hinreichend differenziert expliziert worden? Schließt eine hinreichend differenzierte Explikation die von Kohlberg angegebenen Inhalte (Elemente) aus, und schließt sie nur diese aus bzw. läßt sie die von Kohlberg aufgeführten Inhalte und nur diese zu?

Vor allem aber wurde auf dieser Analyseebene die Frage geklärt, inwieweit Kohlbergs Stufenbestimmungen hinsichtlich des Kriteriums der Konsistenz und Trennschärfe defizitär sind. Inwiefern beruhen zum Beispiel Stufenunterschiede auf Bedeutungsunterstellungen? Dafür steht etwa der Fall von CJs, die explizit - aus der Sichtweise der Subjekte selbst - nur konkrete Rollenerwartungen (Stufe 3) beinhalten, aber im Sinne "gesellschaftlich delegierter" bzw. "gesellschaftlich definierter" Rollenerwartungen interpretiert und damit der für Stufe 4 konstitutiven Systemperspektive zugeordnet werden. Die auf bedeutungsanalytischer Ebene entwickelten Untersuchungsgesichtspunkte finden vor allem in dem Abschnitt "Das Problem der Vermittlung von rationaler und phänomenologischer Rekonstruktion" sowie in dem Abschnitt "Analyse des Standard Issue Scoring-Manuals" Berücksichtigung (S. 49ff.).

### 1.3.3 *Programm der Reanalyse auf empirisch-analytischer Ebene*

Hier wurde eine Kritik an der Analysestrategie von Kohlberg und der an ihm orientierten Forschung vorgenommen:

Die empirischen Analysen Kohlbergs erfolgten sowohl hinsichtlich der Stufen als auch hinsichtlich der Inhaltsaspekte vorrangig auf dem Niveau hoch aggregierter Werte. Niedrig aggregierte Werte wurden stark vernachlässigt, was angesichts des Anspruchs phänomenologischer Rekonstruktion, den Kohlberg erhebt, die Konsequenz hatte, daß die kognitiven Prozesse der Subjekte selbst in ihrer Konkretheit außer acht gelassen wurden.

Auf mikroanalytischer Ebene liegen erste Ergebnisse zur Konsistenz und Kontextbedingtheit der Stufenwerte vor. Dabei zeigte sich eine stärkere Inkonsistenz und Kontextbedingtheit moralischer Urteilsbildung, als Kohlberg annahm. Darüber hinaus wurde das Problem der Interraterreliabilität bei der Kodierung von Criterion Judgments untersucht. Die auf empirisch-analytischer Ebene entwickelten Untersuchungsgesichtspunkte und Resultate finden vor allem im Abschnitt "Ergebnisse auf empirisch-analytischer Ebene" Berücksichtigung (S. 61ff).

Die kritischen Befunde, die auf bedeutungsanalytischer und normativ-analytischer Ebene gewonnen werden, bleiben nicht ohne Folgen für die Bedeutung, d.h. für die konfirmatorische oder falsifikatorische Funktion, die dadurch angeleitete empirische Untersuchungen für bestimmte Hypothesen hinsichtlich der Kontextbedingtheit und Konsistenz haben können. Wenn z.B. nachweisbar ist, daß bei personalen Dilemmata (Joe- und Judy-Dilemma) im Vergleich zu transpersonalen ("Officer-Brown" und Korea-Dilemma) relativ häufig konkrete Rollenerwartungen, wie sie bereits für Stufe 3 kennzeichnend sind, im Sinne einer Systemperspektive interpretiert und damit Stufe 4 zugeordnet werden, so sind in präjudikativer Form erwartbare Kontexteffekte bei den Stufen von vornherein in gewissem Umfang eingegebenet. Eine konfirmatorische oder falsifikatorische Funktion für bestimmte Hypothesen (z.B. die Hypothese kontextbedingter Stufenhäufigkeit) können die Befunde nur im Rahmen einer kohärenten Theorie erfüllen. Deshalb ist es unerläßlich, bestimmte Hypothesen nicht nur auf der Basis der CJs aus Kohlbergs Manual, sondern ebenfalls anhand eines revidierten Manuals zu überprüfen.

Im Hinblick auf weitere empirische Analysen ergeben sich folgende Fragestellungen:

Unter bedeutungsanalytischen Gesichtspunkten sollte untersucht werden, inwieweit Kohlbergs Stufenbestimmungen sowohl auf theoretischer als auch auf operationaler Ebene dem Anspruch rationaler Rekonstruktion gerecht werden und inwieweit es Kohlberg gelingt, die Standards rationaler Rekonstruktion mit dem Anspruch einer phänomenologischen Rekonstruktion der (bedeutungskonstituierenden) Sichtweise der Subjekte selbst zu vermitteln.

Auf empirisch-analytischer Ebene sollte anhand des Kohlberg-Manuals untersucht werden, inwieweit die Variabilität der moralischen Urteilsstufen, von denen die Subjekte Gebrauch machen, kontextbedingt ist. Unter Kontexten sind dabei ausschließlich kognitiv-interne Kontexte zu verstehen, nämlich unterschiedliche Issue-, Frage- und Dilemmatypen (s. Teo, Becker & Edelstein, 1995).

Schließlich stellt sich die Frage, ob bzw. inwiefern sich bestimmte Formen der Inkonsistenz unter dem Rekonstruktionsgesichtspunkt situationsadäquater Urteilsbildung als "kohärente Sinn Ganzheiten" verstehen lassen - ein Gesichtspunkt, der eine genuin moralpsychologische Interpretation strukturierter Ganzheit (a) innerhalb eines Dilemmas und (b) im Vergleich von Dilemmatypen (z.B. personal vs. transpersonal) erlaubt. Dieser hermeneutische Rekonstruktionsgesichtspunkt ist mit gewissen Formen der Inkonsistenz moralischen Urteilens, wie sie auf rein psychometrischer Ebene mittels einer atomistischen Interviewauswertung feststellbar sind, ggf. selbst dann zu vereinbaren, wenn die Urteilsbildung über drei Stufen streut, die Spannweite der Streuung also nicht nur im Sinne des Übergangs von einer Stufe (n) zu einer Stufe (n+1) interpretierbar ist.

Zu (a): Mittels des Konzepts situationsadäquater Urteilsbildung wird innerhalb eines Dilemmas untersucht, welche Funktion einzelne Argumentationsschritte bzw. -stufen in komplexen Argumentationsgängen haben. So kann es sich um die Korrektur einer abstrakt prinzipienorientierten Sichtweise durch eine Beziehungsperspektive handeln, die der besonderen Beziehung Rechnung trägt, in der bestimmte Personen zueinander stehen. Oder es kann sich um eine rekursive Begründung handeln: Pflichten, die eine Person (A) gegenüber einer Person (B) aufgrund der besonderen Beziehung (R) hat, werden im Rekurs auf universelle Grundrechte - z.B. der Gleichbehandlung - relativiert.

Zu (b): Im Rahmen eines Vergleichs unterschiedlicher Dilemmatypen wird anhand der auf dem Kohlberg-Manual und der auf einem revidierten Manual basierenden Stufenwerte die Frage erörtert, inwieweit bei transpersonalen Dilemmata signifikant häufiger von einer

abstrakten Systemperspektive Gebrauch gemacht wird als bei personalen Dilemmata und inwieweit bei personalen Dilemmata häufiger auf konkrete Rollenerwartungen Bezug genommen wird. Dabei ist der manuskritische Befund zu berücksichtigen, daß im Kohlberg-Manual gewisse CJs, die explizit nur konkrete Rollenerwartungen beinhalten und die in personalen Dilemmata relativ häufig auftreten, als "gesellschaftlich definierte" oder "gesellschaftlich delegierte" interpretiert und damit im Sinne einer Systemperspektive umgedeutet werden. In Fällen, in denen postkonventionelle mit konventionellen Argumentationsformen verknüpft sind, wird untersucht, ob bzw. inwieweit postkonventionelle Sichtweisen bei personalen Dilemmata signifikant häufiger mit einer Beziehungsperspektive und bei transpersonalen Dilemmata häufiger mit einer abstrakten Systemperspektive verknüpft sind.

Das Konzept situationsadäquater Urteilsbildung ist ein relativ komplexes normatives Konzept. Empirisch ist lediglich eine mehr oder weniger hohe Inkonsistenz moralischen Urteilens, die sich eventuell als kontextbedingt erweist, feststellbar. Das bloß Feststellbare wird erst unter dem Rekonstruktionsgesichtspunkt situationsadäquater Urteilsbildung aus kognitiv-internen Gründen verstehbar. Um die Rekonstruktion des aus kognitiv-internen Gründen Verstehbaren aber geht es primär im Rahmen eines kognitiv-strukturtheoretischen Forschungsansatzes. An dieser Stelle muß wenigstens erwähnt werden, daß sich das Konzept situationsadäquater Urteilsbildung auch unter differentiellen Gesichtspunkten als fruchtbar erweist; es erlaubt nämlich die Differenzierung zwischen einer Personengruppe, die sich durch den flexiblen Gebrauch von unterschiedlichen Argumentationsformen bzw. Argumentationsstufen im Sinne situationsadäquaten Urteilens kennzeichnen läßt, und einer Personengruppe, bei der dies nicht der Fall ist.

## **2. Ergebnisse auf theoretisch-konzeptueller Ebene**

### 2.1 Das Problem der Vermittlung von rationaler und phänomenologischer Rekonstruktion

Die Rekonstruktion der Entwicklung moralischer Urteilsfähigkeit erfolgt bei Kohlberg zunächst von einer phänomenologischen Perspektive aus; es geht dabei um die Erfassung der bedeutungskonstituierenden Sichtweise der Subjekte selbst: "The researcher must operate phenomenologically by looking at morality from the subject's viewpoint, understanding what the subject is saying in his or her own terms. That is, the researcher is seeking to understand what the subject means when making moral judgments rather than attributing meaning to the

judgments from some outside system of interpretation not shared by the subject" (Colby & Kohlberg, 1987, S. 1ff.).

Voraussetzung für eine rationale Rekonstruktion der Entwicklung moralischer Urteilsfähigkeit sind Stufenbestimmungen, die (a) hinreichend konsistent und trennscharf sind und (b) die Kriterien einer universellen Entwicklungslogik erfüllen. Die Rekonstruktion dieser Entwicklungslogik erfolgt dabei im Hinblick bzw. im Vorgriff auf einen ausgezeichneten normativen Bezugspunkt, der sich bei Kohlberg im Sinne des "ideal role taking", also im Sinne einer vollständig reversibel gewordenen Perspektivenübernahme definiert: "Our theory is a rational reconstruction because it (a) describes the developmental logic inherent in the development of justice reasoning with the aid of (b) the normative criterion of stage 6 which is held to be the most adequate (i.e. most reversible) stage of justice reasoning" (Kohlberg 1984, S. 221).

Der Anspruch einer rationalen Rekonstruktion, an dem Kohlbergs Stufenmodell moralischer Urteilsfähigkeit als Kompetenztheorie seinen grundlegenden Maßstab hat, ist zwar mit dem Anspruch einer phänomenologischen Rekonstruktion der (bedeutungskonstituierenden) Sichtweise der Subjekte selbst untrennbar verknüpft, steht aber zugleich in einem eigenartigen Spannungsverhältnis dazu. Nichts ist im Rahmen einer Kohlberg-Rekonstruktion schwerer auf den Punkt zu bringen als dieses Spannungsverhältnis. Der Versuch, es reduktionistisch zugunsten rein rationaler Rekonstruktion (im Sinne einer "normativen Theorie guter Gründe" auf deduktivem Wege) oder einer rein phänomenologischen Rekonstruktion (im Sinne einer empirischen Theorie auf dem Wege induktiver Generalisierung) aufzulösen, liefe auf die Preisgabe von Kohlbergs Forschungsansatz hinaus.

Es war zu untersuchen, inwieweit Kohlbergs Stufenmodell moralischer Urteilsfähigkeit den Anspruch einer sowohl rationalen als auch phänomenologischen Rekonstruktion erfüllt.

Im folgenden soll zunächst dargestellt werden, an welchen Kriterien der Anspruch rationaler Rekonstruktion auf den Ebenen (a) theoretischer und (b) operationaler Stufenbestimmungen bei Kohlberg festgemacht ist.

Zu (a): Auf theoretischer Ebene umfaßt dieser Rekonstruktionsanspruch 6 Stufen. Dabei stellen die Stufen 1 bis 5 Stufen natürlicher Argumentation dar, die empirisch repräsentiert sind. Stufe 6 bildet den ausgezeichneten normativen Bezugspunkt rationaler Rekonstruktion, der zwar empirisch nicht repräsentiert ist, in dessen Lichte aber die Rekonstruktion aller anderen Stufen erst erfolgen kann (vgl. Habermas, 1983). Der normative Bezugspunkt rationaler



Rekonstruktion definiert sich durch den Standpunkt einer autonomen und universalistischen Moral in der Tradition rationaler Normenbegründung. Dieser Standpunkt ist durch vier moral-philosophische Traditionen repräsentiert: durch das Modell des Gesellschaftsvertrags (Rawls); durch die kantianisch-deontische Tradition, die an der Achtung vor der Würde von Personen und der Universalisierbarkeit ihrer Handlungsmaximen orientiert ist; durch eine utilitaristische Tradition, die an den Folgen einer Handlungsweise für die Mehrung des allgemeinen Wohls ihren Maßstab hat; sowie durch die Diskursethik, welche die Universalisierbarkeit moralischer Normen an die universalpragmatischen Bedingungen rationaler Argumentation bindet (Apel, 1988; Habermas, 1991).

Die rationale Rekonstruktion der Stufen 1 bis 6 nimmt Kohlberg auf der Grundlage von sozio-moralischen Perspektiven vor, durch die Verstehensleistungen ermöglicht werden, die im Sinne eines irreversiblen Voraussetzungs-zusammenhangs aufeinander bezogen sind. Die Perspektivenübernahme einer Stufe (n+1) beruht dabei im Vergleich zu einer Stufe (n) auf einer komplexeren Differenzierung und Koordination konfligierender moralischer Sichtweisen; sie zeichnet sich durch das höhere Rationalitätspotential aus, das für moralische Entscheidungs- sowie Normbegründungen zur Verfügung steht und das es erlaubt, bestimmte Begründungsformen als (mehr oder weniger) intersubjektiv verbindlich bzw. universalisierbar zu verstehen: "With regard to the question of each stage presupposing the understanding achieved at lower stages, the level of perspective again provides a convincing rationale. The claim is simply that one cannot take a prior-to-society perspective without the capacity to think in terms of a social system; one cannot take a social system perspective without the capacity to think in terms of shared norms and expectations of a group or mutually oriented dyadic relationships; one cannot respond in terms of coordinated individual perspectives until one can differentiate individual perspectives" (Colby et al., 1983, S. 66).

Zu (b): Auf der durch die Criterion Judgments repräsentierten, operationalen Ebene der Stufenbestimmungen definieren sich Stufen durch die Verknüpfung von Struktur- und Inhaltsaspekten des moralischen Urteilens. Der Anspruch einer rationalen Rekonstruktion der Stufensequenz moralischer Urteilsfähigkeit bezieht sich deshalb auf operationaler Ebene nicht nur auf die Strukturaspekte, sondern auch auf die Inhaltsaspekte bzw. auf bestimmte Struktur-Inhaltsrelationen.

Kohlberg erfaßt bei Criterion Judgments, die durch die Kombination gleicher Normen und Elemente gekennzeichnet sind, Inklusionsbeziehungen: "Within a single dilemma and issue, a set of hierarchical inclusions has been defined so that a particular CJ may be considered to include within it other CJs and/or to be included itself within other CJs. A given CJ may

include lower stage CJs that express parallel but developmentally less advanced concepts. Alternatively, a given CJ may be included within higher stage CJs that express parallel but developmentally more advanced concepts.

In order for one CJ to be defined as included within another, the idea expressed by the higher stage CJ must logically entail acceptance of the simpler or more partial idea expressed by the lower stage CJ. These inclusion relationship most often but not always involve CJs with the same norm and element but different stages" (Colby & Kohlberg, 1987, S. 177).

Im Gegensatz zu den theoretischen Stufenbestimmungen erfolgen die operationalen Stufenbestimmungen nur im Hinblick auf die Stufen 1 bis 5, schließen aber auch Zwischenstufen ein.

Unseres Erachtens sind sowohl Kohlbergs theoretische wie auch seine operationalen Stufenbestimmungen unter dem Gesichtspunkt rationaler Rekonstruktion partiell problematisch und revisionsbedürftig. Wir nehmen zunächst eine Neubestimmung postkonventionellen Denkens bzw. der Differenzierung zwischen postkonventioneller und konventioneller Argumentation vor (2.1.1), entwickeln darüber hinaus aber auch eine kritische Rekonstruktion des konventionellen und des präkonventionellen Niveaus moralischer Urteile (2.1.2).

Was die operationalen Stufendefinitionen angeht, so ist unter empirisch-analytischen Gesichtspunkten bei Kohlbergs Manualkonstruktion die schmale Basis von nur 7 - nach dem Zufallsprinzip ausgewählten - Konstruktionsfällen problematisch. Da auf dieser Basis die CJs, d.h. die operationalen Stufenbestimmungen, gewonnen werden, durch die im voraus festgelegt wird, welche "möglichen Sachverhalte" empirisch überhaupt erfaßbar sind, ist bei der Kodierung von vornherein eine mehr oder weniger weitreichende Ausblendung bestimmter Interviewäußerungen bzw. Argumentationsmuster zu befürchten. Es bestätigt sich, daß es aufgrund der schmalen empirisch-analytischen Konstruktionsbasis des Manuals gelegentlich zu einer minimalen Interviewauswertung kommt. Vor allem auf präkonventionellem und postkonventionellem Niveau kommt es in größerem Ausmaß zu dieser minimalen Interviewauswertung.

### *2.1.1 Das Problem einer angemessenen Differenzierung zwischen konventionellen und postkonventionellen moralischen Urteilen*

Unter normativ- und begriffsanalytischen Gesichtspunkten muß auf der Ebene der operationalen Stufenbestimmungen, d.h. auf CJ-Ebene, zunächst auf das Problem einer - im Vergleich zu Kohlberg - angemesseneren Abgrenzung der postkonventionellen gegenüber der konventionellen Moral eingegangen werden.

Wir haben, um unter begriffs- und normativ-analytischen Gesichtspunkten zwischen den Stufen konventioneller Moral, der Übergangsstufe 4/5 und der postkonventionellen Stufe 5 in hinreichend trennscharfer, angemessener Weise unterscheiden zu können, ein Klassifikationssystem eingeführt, das auf der Trennung und Kombination von (a) "universalistischen" vs. "nicht-universalistischen" und (b) "autonomen" vs. "nicht-autonomen" Begründungsformen beruht.

Im Rahmen des von uns entwickelten Klassifikationssystems werden neben dem Standpunkt einer "autonomen und universalistischen Moral im emphatischen Sinne" (Stufe 5 und 6) zusätzliche Differenzierungen vorgenommen: (a) zwischen "einer autonomen Moral nicht-universalistischer Art" und (b) einer "universalistischen Moral nicht-autonomer Art". Außerdem wurde (c) noch zwischen einer "autonomen Moral im starken und im schwachen Sinne" differenziert.

Im Rahmen dieses Klassifikationssystems definiert sich die postkonventionelle Stufe 5 durch eine soziomoralische Perspektivenübernahme, die vom Standpunkt einer autonomen und universalistischen Moral im starken Sinne erfolgt. Die Übergangsstufe 4/5 definiert sich hinsichtlich einer Perspektivenübernahme, die auf dem Standpunkt einer

- autonomen und universalistischen Moral im schwachen Sinne oder einer
  - autonomen und nicht-universalistischen Moral oder
  - einer universalistischen und nicht-autonomen Moral
- beruht.

Das konventionelle Niveau moralischer Urteilsfähigkeit hingegen ist per negationem dadurch definiert, daß die soziomoralische Perspektivenübernahme von keinem der o.g. Standpunkte der Moral aus erfolgt.

Der Begriff "autonome Moral" hat eine zweifache Bedeutung. Zum einen ist darunter eine "subjektive Autonomie" in engeren moralischen Sinne zu verstehen. In Berufung auf individuelle Gewissensentscheidungen oder innere Überzeugungen setzt ein isoliertes moralisches Selbst seine eigenen, selbstgesetzten Standards des moralisch "Richtigen" sozial etablierten Standards und Legitimationsinstanzen entgegen: Vgl. etwa Heinz-Dilemma, Issue: "Leben", CJ #33, Stufe 4/5: "[Heinz should steal the drug] if he considers the value of the individual's right to life as greater than that of established laws of society; or because although by the standards of society stealing is wrong, by Heinz's standards as an individual, stealing to save a life could be morally right."

Diese Bedeutungskomponente "autonomer Moral" schließt keinen Rekurs auf die moralische Substanz universeller Grundrechte ein und kann deshalb nicht der Stufe 5, sondern nur der Übergangsstufe 4/5 zugeordnet werden. In dieser Hinsicht muß keine Korrektur an Kohlbergs Einstufungen vorgenommen werden.

Zum anderen ist unter "autonomer Moral" ein moralischer Standpunkt zu verstehen, der (a) auf einer Konzeption der "Eigengesetzlichkeit" der Moral selbst beruht und zwar in dem Sinne, daß es zu einer klaren Ausdifferenzierung der unterschiedlichen Wert- und Geltungssphären von Religion, Recht, Moral, Konvention usf. in ihrer Eigensinnigkeit kommt, eine Konzeption, die zugleich (b) den Rekurs auf universelle Personrechte und universelle Verfahren legitimer Normsetzung und -durchsetzung einschließt. Von einer "autonomen Moral universalistischer Art" in diesem emphatischen Sinne kann nur dann die Rede sein, wenn alle überhaupt denkbaren Grenzen partikularer Gruppenmoral transzendiert werden, d.h. wenn es zu keiner Ausgrenzung bestimmter sozialer Gruppen oder Lebens- und Denkformen kommt, es sei denn, diese stünden in Widerspruch zur moralischen Substanz der Menschenrechte. In Übereinstimmung mit Kohlberg ordnen wir nur diese Begründungsform der postkonventionellen Stufe 5 zu.

Im Gegensatz zu diesem emphatischen Begriff einer autonomen universalistischen Moral, der auf einer klaren Ausdifferenzierung von unterschiedlichen Wert- und Geltungssphären im o.g. Sinne sowie dem expliziten Rekurs auf universelle Grundrechte und Legitimationsverfahren beruht, kann von einer autonomen universalistischen Moral auch in einem schwächeren und relativ konkretistischen Sinne die Rede sein (vgl. das "Gleichnis vom barmherzigen Samariter!"): alle "menschlichen Wesen" haben, sofern Gefahr für Leib und Leben droht und sie (unverschuldet) in Not geraten sind, gleichermaßen Anspruch auf Hilfeleistung bzw. auf Rettung - vgl. z.B. Heinz-Dilemma, Issue "Leben", Stufe 3, CJ#9: "[Heinz should steal the drug even he doesn't love his wife, or even for a stranger] because we are all human beings and should be willing to help others; or because it would be inhuman not to save her just because he doesn't love her; or because his wife is still a human being, or a human life is always precious." - Es scheint sinnvoll, auch "universalistische Begründungen, die nur im schwachen Sinne autonomer Art" sind, auf der Übergangsstufe 4/5 anzusiedeln (s. das bereits zitierte CJ#9, Heinz-Dilemma, Issue: "Leben", das bei Kohlberg auf dem Wege einer problematischen Vereindeutigung und Festlegung im voraus auf Stufe 3 angesiedelt wird). Diese nur im schwachen Sinne "universalistischen Begründungen autonomer Art" sind per negationem dadurch gekennzeichnet, daß sie weder eine Bezugnahme auf metaphysisch-religiöse Welt- und Selbstinterpretationen noch einen Rekurs auf die "Eigengesetzlichkeit"

der Moral, d.h. eine klare Ausdifferenzierung der unterschiedlichen Wert- und Geltungssphären von Religion, Recht, Moral und Konvention einschließen. Universalistische Begründungen, die nur im schwachen Sinne autonomer Art sind, erfolgen in Bezugnahme auf die Goldene Regel oder in stärker alterorientierter empathischer Einstellung vom Standpunkt einer Liebes- und Mitleidsethik aus, die sich geschwisterlich allen Menschen und darüber hinaus allen lebendigen Kreaturen als verletzbar, bedürftigen Wesen verbunden weiß. Nicht die rationale, sondern die emotivistische Komponente einer autonomen universalistischen Moral ist insofern vorherrschend.

Wenn zwar einerseits bereits universalistische Sichtweisen und Standards entwickelt werden, die - der Intention nach - die Funktion haben, die Grenzen partikularer Gruppenmoral und damit die Grenzen rein konventioneller Moral zu transzendieren, diese Universalisierungen aber andererseits noch in den Kontext partikularer Lebens- und Denkformen eingebettet bleiben, die eine metaphysisch-religiös fundierte Legitimationspraxis haben (vgl. Ben-Chorin, 1977; Roetz, 1992; Vasudev, 1986), so handelt es sich um "universalistische Begründungsformen nicht-autonomer Art": vgl. z.B. Heinz-Dilemma, Issue: "Leben", Stufe 4, CJ#29: "[Heinz should steal the drug] ... because saving a human life derives from fundamental religious laws which man-made laws may add to but cannot displace."

Alle Weltreligionen haben Formen einer universalistischen Ethik entwickelt, die die Grenzen partikularer Gruppenmoral und damit auch die Grenzen einer konventionellen Moral in gewisser Weise transzendieren. Aus den genannten Gründen wäre es jedoch unangebracht, bereits universalistische Begründungsformen nicht-autonomer Art als postkonventionelle moralische Denkformen zu begreifen. Sinnvollerweise können universalistische Begründungsformen, die nicht vom Standpunkt einer autonomen Moral aus entwickelt werden, nur als Übergangsformen von konventioneller zu postkonventioneller Moral verstanden werden.

Im Rahmen kulturvergleichender Forschung (Boyes & Walker, 1988) erweist sich Kohlbergs Auswertungsmanual hinsichtlich universalistischer Begründungsformen nicht-autonomer Art in zweifacher Weise als defizitär: zum einen in der Hinsicht, daß diese Begründungsformen in Kohlbergs Auswertungssystem unzureichend repräsentiert sind; zum anderen hinsichtlich ihrer Zuordnung. Derartige Begründungsformen werden bei Kohlberg nämlich unangemessen der Stufe 4 auf dem Niveau konventioneller Moral zugeordnet. Mit Nachdruck muß freilich hervorgehoben werden, daß Kohlberg postkonventionelles Denken (im Sinne der natürlichen Argumentationsstufe 5) zu Recht an den Standpunkt einer autonomen und universalistischen Moral im emphatischen Sinne gebunden hat. "Universalistische Begründungsformen nicht-

autonomer Art " sind dadurch gekennzeichnet, daß sie noch an unterschiedliche Formen soziokultureller Identität, für die unterschiedliche symbolische Repräsentationsformen und metaphysisch-religiöse Kodiersysteme der Welt- und Selbstinterpretation konstitutiv sind, gebunden bleiben. Da diese Bindung im Kontext von universalistischen Sichtweisen nicht-autonomer Art nicht mehr zum Gegenstand der Reflexion bzw. kritischer Distanzierung gemacht werden kann, stellen derartige Sichtweisen selber eine Quelle dramatischer Konflikte dar, die nur im Rahmen einer "autonomen und universalistischen Moral im emphatischen Sinne" vermieden werden können oder eine für alle Konfliktparteien gleichermaßen akzeptable Konfliktlösung finden können.

Der Standpunkt einer autonomen und universalistischen Moral im emphatischen Sinne stellt zwar hinsichtlich seiner historischen Konstitutionsbedingungen ein spezifisches Produkt europäischer Aufklärung und Reflexionskultur dar und ist insofern als kulturrelativ zu verstehen, aber damit ist sein universeller Geltungsanspruch als der denkbar höchste Standpunkt der Moral, zu dem es keine kulturrelativistisch verstehbare Alternative gibt, noch überhaupt nicht tangiert. Dieser universelle Geltungsanspruch beruht ausschließlich auf seiner überlegenen Problemlösungskapazität, d.h. auf der uneingeschränkten Reversibilität der Perspektivenübernahme, die dieser Standpunkt der Moral ermöglicht.

Der (subjektivistische) Standpunkt einer "autonomen Moral nicht-universalistischer Art" ist von Kohlberg zu Recht Stufe 4/5 zugeordnet worden. Im Unterschied zu Kohlberg sind aber auch Begründungsformen, die nur im "schwachen Sinne autonomer und universalistischer Art" sind (Kohlbergstufe 3) sowie "universalistische Begründungen nicht-autonomer Art" (Kohlbergstufe 4) auf der Übergangsstufe 4/5 anzusiedeln.

Kohlbergs Kennzeichnung der Stufe 5 ist nur auf der Ebene theoretischer Stufenbestimmungen unproblematisch. Auf der operationalen Ebene der CJs hingegen bleibt z.T. unklar, ob bestimmte "Criterion Judgments" und "Match Examples" zurecht als prototypische Äußerungen im (strukturellen) Sinne postkonventioneller Moral verstanden werden können. Unklar ist dies vor allem in den Fällen, in denen - in relativ plakativer Form - lediglich auf isolierte, einzelne Menschenrechte oder pauschal auf Grundrechte Bezug genommen wird: Heinz-Dilemma, Issue: "Gesetz", Stufe 5, CJ#33: "[One should obey the law] insofar as the law protects the basic rights of some individuals against infringement of others."

Joe-Dilemma, Issue: "Vertrag", Stufe 5, CJ#39: "[Joe should refuse to give his father the money, or property should be respected] because Joe is an individual whose personal rights

are of equal value to his father's; or because each person as a free individual is entitled to enjoy the use of property as long as it does not infringe upon the like rights of others."

Eine solche (inhaltliche) Bezugnahme ist im Rahmen demokratischer und rechtsstaatlich verfaßter Gesellschaften, insbesondere in den USA, bei allen sozialen Gruppen, die eine positive Berufs- und Lebensperspektive haben, wahrscheinlich zum Bestandteil konventioneller Moral geworden.

Um hinreichend klar zwischen konventionellem und postkonventionellem moralischem Denken unterscheiden zu können, bedarf es eines Interviewleitfadens, der darauf zielt, ob und inwieweit die Kompetenz, konzeptuelle Differenzierungen im Sinne unseres Klassifikationssystems vorzunehmen, entwickelt ist.

In letzter Zeit ist das Problem der Ausdünnung postkonventioneller Moral, zu dem Kohlbergs Stufenmodell aufgrund eines kulturspezifisch verengten oder zu hoch angesetzten Konzepts postkonventioneller moralischer Urteilsfähigkeit angeblich geführt hat, intensiver diskutiert worden (vgl. Eckensberger, 1995; Shweder et al., 1987; Snarey, 1985). Diese Diskussion bzw. die Kohlbergkritik, die geübt wird, kann aber nicht überzeugen, da sie keine unter normativ-analytischen Gesichtspunkten hinreichend abgeklärte Grundlage hat, ein Mangel, dem wir durch die konzeptuellen Differenzierungen, wie sie in dem von uns eingeführten Klassifikationssystem entwickelt werden, abzuhelpen versuchen.

Wird Kohlbergs Bestimmung postkonventioneller Moral und die (strukturelle) Differenzierung, die im Kohlberg-Manual zwischen konventionellen und postkonventionellen moralischen Urteilen erfolgt ist, an den Standards unseres Klassifikationssystems gemessen, so gelangt man hinsichtlich des Problems der Ausdünnung postkonventioneller Moral zu folgenden Schlußfolgerungen: (a) es ist sinnvoll, an Kohlbergs theoretischen Stufenbestimmungen postkonventioneller Moral (Stufe 5 und Stufe 6) festzuhalten, prinzipiell jedenfalls. (b) Daraus folgt, daß nicht hinsichtlich Stufe 5, sondern lediglich in Beziehung auf die Übergangsstufe 4/5 von einer konzeptuell bedingten Ausdünnung sinnvollerweise die Rede sein kann.

Die klassifikatorische Differenzierung zwischen (a) universalistischen Begründungsformen autonomer vs. nicht-autonomer Art und (b) autonomen Begründungsformen universalistischer vs. nicht-universalistischer Art hat zwar gegenüber den unter normativ-analytischen Gesichtspunkten nicht hinreichend ausgewiesenen, z.T. problematischen Stufenzuordnungen bei Kohlberg den Vorzug, daß sie es erlaubt, zwischen (rein) konventioneller Moral und

postkonventioneller Moral sowie bestimmten Übergangsformen klarer zu unterscheiden und im Sinne vollständiger Disjunktion alle denkbaren Fälle zu erfassen und auszuschöpfen. Dieser Vorzug kann jedoch nur dann in angemessener Weise zur Geltung gebracht werden, wenn die genannten Begriffsbestimmungen und Kriterien bei den konkreten Stufenbestimmungen im Auswertungsmanual hinreichende Berücksichtigung finden: Konkret bedeutet das, daß das Manual einer entsprechenden Revision zu unterziehen ist.

Im Rahmen einer rationalen Rekonstruktion moralischer Urteilsfähigkeit, die im Lichte einer autonomen und universalistischen Moral erfolgt, ist es unter normativ-analytischen Gesichtspunkten einsichtig, warum eine universalistische Begründungsform autonomer Art gegenüber einer universalistischen Begründungsform nicht-autonomer Art den Vorzug verdient und in diesem Sinne als höherstufig zu verstehen ist.

Die Standards rationaler Rekonstruktion erlauben es, die Entwicklungssequenz der Stufen x, y, z (z.B. der Stufen 4, 4/5 und 5) im Sinne normativer Präferenzrelationen zu verstehen, die aus rein kognitiv-internen, d.h. rationalen Gründen einsichtig sind. Rationalitätsstandards erklären nicht, warum unter bestimmten Entwicklungsbedingungen mikrosozialer oder makrosozialer Art (die sich kulturtheoretisch durch soziokulturelle Lebensformen und symbolische Repräsentations- und Kodierungsformen der Welt- und Selbstinterpretation definieren) mit relativ hoher Wahrscheinlichkeit eher diese als jene Entwicklungsmuster moralischer Urteilsbildung auftreten und ob bzw. inwiefern sich die Erwartbarkeit bestimmter Entwicklungsmuster als kontextabhängig, altersabhängig oder personabhängig erweist. Eine Klärung entsprechender Fragestellungen kann nur auf empirisch-analytischer Ebene erfolgen.

Auch folgende grundlegende Fragen sind nur auf empirisch-analytischer Ebene entscheidbar: Unter welchen Entwicklungsbedingungen, die sich unter mikro- oder makrosozialen Gesichtspunkten als mehr oder weniger entwicklungsförderlich bzw. restriktiv erweisen, erfolgen die Entwicklungsverläufe mit deutlich überzufälliger Erwartbarkeit in der Weise, daß erst einmal eine Begründungsform x (z.B. eine "universalistische Sichtweise nicht-autonomer Art") entwickelt wird, bevor y (z.B. eine "universalistische Sichtweise autonomer Art") entwickelt werden kann.

Unter welchen spezifischen Entwicklungsbedingungen ist eine bestimmte (Stufen-)Sequenz von Begründungsformen, die durchlaufen wird, als ein irreversibler Voraussetzungs Zusammenhang soziomoralischer Verstehensleistungen zu interpretieren? Dieser irreversible Voraussetzungs Zusammenhang würde sich indessen aus bestimmten externen Bedingungen (zu denen es im Rahmen bestimmter Lebensformen keine Alternative gibt) erklären und im



bloßen Rekurs auf kognitiv-interne, d.h. rationale Entwicklungsvoraussetzungen, nicht nur schlecht, sondern überhaupt nicht begreifen lassen. In Abhängigkeit von welchen externen Entwicklungsbedingungen treten dagegen mit relativ hoher Erwartbarkeit unterschiedliche Entwicklungspfade auf?

Auf empirisch-analytischer Ebene stellt sich die Frage, ob die Begründungsform  $x$  im Zuge der Entwicklung moralischer Urteilsfähigkeit übersprungen werden kann und unter welchen Entwicklungsbedingungen dies der Fall ist. Auch die komplementäre Frage, ob und inwiefern es unter bestimmten Entwicklungsbedingungen zu Regressionen kommt, ist nur auf empirisch-analytischer Ebene zu klären ("Regressionen" definieren sich dabei hinsichtlich der o.g. normativ-analytischen Gesichtspunkte). Es kann auch nur auf dieser Ebene untersucht werden, unter welchen Bedingungen zu erwarten ist, daß Entwicklungsverläufe auftreten, die durch ein bestimmtes (kombinatorisches) Muster von Überspringen und Regressionen gekennzeichnet sind. Es ist evident, daß derartige Untersuchungen im Rahmen kulturvergleichender Forschung in hohem Maße relevant und wünschenswert wären. Folgende alternative Entwicklungspfade hinsichtlich des Übergangs von konventioneller zu postkonventioneller Moral sind denkbar, d.h. sie können unter begriffsanalytischen und normativ-analytischen Gesichtspunkten als "mögliche Sachverhalte" konstruiert werden, die theoretisch zugelassen und unter moralpsychologischen Gesichtspunkten plausibel sind:

- (a) die etablierten Standards partikularer Gruppenmoral werden zunächst einmal im Sinne von Begründungsformen universalistischer, aber nicht-autonomer Art transzendiert. In einem weiteren Entwicklungsschritt wird dann der Standpunkt nicht-autonomer Moral überwunden: es wird eine universalistische Sichtweise autonomer Art entwickelt, und zwar ( $\alpha$ ) im schwachen oder ( $\beta$ ) im strengen, emphatischen Sinne. Dabei besteht zwischen ( $\alpha$ ) und ( $\beta$ ) eine starke voraussetzungslogische Relation; ( $\alpha$ ) tritt in jedem Falle vor ( $\beta$ ) auf.

- (b) Der Entwicklungsfortschritt vollzieht sich in der Weise, daß die Entwicklung universalistischer Begründungen, die nicht-autonomer Art sind, "übersprungen" wird, d.h. von konventioneller Moral wird unmittelbar zu ( $\alpha$ ) einer autonomen Moral nicht-universalistischer Art oder ( $\beta$ ) zu einer autonomen Moral universalistischer Art fortgeschritten.

Auch der Zwischenschritt einer autonomen Moral nicht-universalistischer Art ist nicht erforderlich; auch ( $\alpha$ ) kann übersprungen werden.

In jedem Falle geht jedoch einer "autonomen, universalistischen Moral" im engeren Sinne eine im schwachen Sinne voraus. Insofern und nur insofern kann beim Übergang von

konventioneller Moral zur postkonventionellen Moral, wie sie durch Stufe 5 repräsentiert ist, Stufe 4/5 nicht übersprungen werden.

- (c) Auch dieser Voraussetzungszusammenhang kann sich als soziokulturell bedingt erweisen, d.h. auch der Standpunkt einer "autonomen und universalistischen Moral im schwachen Sinne" kann übersprungen werden. In diesem Falle wird von konventioneller Moral unmittelbar zu postkonventioneller Moral im Sinne von Stufe 5 fortgeschritten, z.B. auf dem Wege des "Ausbruchs" und der Bewältigung einer tiefgreifenden Orientierungskrise.

Der Rekonstruktionsanspruch, unter dem Kohlbergs konzeptuell angeleitete Manualkonstruktion steht, sowie die systematische und hoch differenzierte Form der Durchführung sind nicht zu überbieten, prinzipiell jedenfalls. Andererseits ist Kohlbergs Manualkonstruktion aufgrund dieses hohen Anspruchs in doppelter Weise angreifbar: unter normativ-analytischen, aber auch unter empirisch-analytischen Gesichtspunkten.

Der von Kohlberg - im starken voraussetzungslogischen Sinne - interpretierte Voraussetzungszusammenhang zwischen angeblich rein konventionellen und postkonventionellen Formen universalistischer Argumentation ist fragwürdig.

Kohlberg versteht den konzeptuell im Manual vorentworfenen Voraussetzungszusammenhang zwischen den Begründungsmustern x und y in dem starken voraussetzungslogischen Sinne, daß in jedem Falle aus kognitiv-internen, rationalen Gründen das Begründungsmuster x dem Muster y vorausgeht, d.h. früher auftritt. Ein Überspringen von x wird theoretisch ausgeschlossen. Sofern dies vorkommt, kann dies nur als Meßfehler verstanden bzw. toleriert werden, der ggf. dadurch zu korrigieren wäre, daß der Abstand der Meßzeitpunkte verkürzt wird. Kohlberg räumt auf theoretischer Ebene lediglich die Existenz von Regressionen ein (Colby et al., 1983), allerdings ohne zu präzisieren, wieviel Stufen Regressionen umfassen dürfen, um noch theoriekompatibel zu sein. Regressionen erklären sich in jedem Falle aus kognitiv-externen Bedingungen moralischer Urteilsbildung; bei Kohlberg wird dabei ausschließlich an "extreme Umstände" gedacht.

Im Gegensatz zu Kohlberg erscheint es uns plausibler, von einem im schwachen voraussetzungslogischen Sinne zu verstehenden Voraussetzungszusammenhang zwischen universalistischen Argumentationen nicht-autonomer vs. autonomer Art auszugehen. Diese voraussetzungslogische Konzeption wäre nur dann widerlegt, wenn eine Klasse von Fällen nachweisbar wäre, in denen autonome Begründungsmuster zeitlich vor nicht-autonomen auftreten. Eingeräumt wird hingegen im Rahmen einer schwachen Voraussetzungslogik, daß

(a) universalistische Argumentationen autonomer und nicht-autonomer Art sowie (b) moralische Gesichtspunkte konventioneller Art und Gesichtspunkte, die der Intention nach die Funktion haben, die Grenzen partikularer Gruppenmoral zu transzendieren, gleichzeitig entwickelt werden können.

Zudem wird die Möglichkeit des Überspringens von universalistischen Begründungsformen nicht-autonomer Art sowie Regressionen von Begründungen autonomer Art auf Begründungen nicht-autonomer Art eingeräumt. Inwieweit bestimmte Verlaufsmuster der Entwicklung moralischen Urteilens auftreten bzw. mit welcher Wahrscheinlichkeit sie erwartbar sind, hängt von den für bestimmte soziokulturelle Lebensformen kennzeichnenden makro- und mikrosozialen Entwicklungsbedingungen in Verknüpfung mit bestimmten symbolischen Repräsentationsformen sowie Kodiersystemen der Welt- und Selbstinterpretation ab.

Im Rahmen eines "sozialkonstruktivistischen Forschungsansatzes", wie er in unterschiedlicher Form von Edelstein (1992) und von Lempert (1988) entwickelt worden ist, empfiehlt es sich unseres Erachtens, Entwicklungssequenzen - zumindest hinsichtlich konventioneller und postkonventioneller Begründungsformen - in einem schwachen voraussetzungslogischen Sinne zu interpretieren und darüber hinaus alternativen Entwicklungspfaden Rechnung zu tragen.

### *2.1.2 Das Problem angemessener Stufenbestimmungen auf konventionellem und präkonventionellem Niveau*

Die Standards rationaler Rekonstruktion ermöglichen keineswegs nur eine angemessene Unterscheidung zwischen rein konventionellen und postkonventionellen Begründungsformen sowie Übergangsformen (im Sinne der Stufe 4/5). Vielmehr erlauben sie es, hinsichtlich aller Stufen zwischen mehr oder weniger tragfähigen bzw. intersubjektiv verbindlichen Begründungsformen zu unterscheiden. Auf allen Stufen ist dabei die Frage zu klären, inwieweit der Anspruch einer rationalen Rekonstruktion stufenspezifischer soziomoralischer Perspektiven mit dem Anspruch einer phänomenologischen Rekonstruktion der (bedeutungskonstituierenden) Sichtweise der Subjekte selbst in Einklang gebracht werden konnte, eine Frage, die sich primär hinsichtlich der Kohlbergschen Manualrekonstruktion unter normativ-analytischen und begriffs- bzw. bedeutungsanalytischen Gesichtspunkten stellt, und zwar vor allem in Beziehung auf zwei Problembereiche: hinsichtlich (a) von Bedeutungsunterstellungen und (b) von Bedeutungsverschiebungen.

"Bedeutungsunterstellung" bezieht sich auf das problematische Verhältnis von CJs und den im Kohlberg-Manual damit verknüpften Stufenstrukturen (stage structures). Die Analyse von "Bedeutungsverschiebungen" hingegen bezieht sich auf Formen logisch-semantischer Inkonsistenz, auf denen die im Kohlberg-Manual vorgenommenen Zuordnungen von bestimmten Bedeutungskomponenten zu bestimmten Stufenstrukturen partiell beruhen, und zwar in Abhängigkeit von Dilemmatypen (personal vs. transpersonal) und issues (autonom vs. heteronom).

Von "Bedeutungsunterstellungen" ist sinnvollerweise dann die Rede, wenn unterschiedliche Interpretationsmöglichkeiten, zwischen denen eine Entscheidung nur (a) im Kontext von Einzelfallanalysen nicht-atomistischer Art oder (b) in Konfrontation mit der Sichtweise und Bewertung der Sichtweise der Subjekte selbst herbeigeführt werden kann, von vornherein im Sinne einer problematischen Vereindeutigung auf eine einzige Interpretationsmöglichkeit reduziert werden.

"Bedeutungsverschiebung" definiert sich dadurch, daß in dilemma- oder issueabhängiger Weise entweder (a) gleiche Strukturaspekte verschiedenen Stufen zugeordnet sind oder (b) unterschiedliche Strukturaspekte gleiche Stufen repräsentieren.

Zu (a): Die Zuordnung gleicher Strukturaspekte zu unterschiedlichen Stufen stellt eine grobe Inkonsistenz dar, die in jedem Falle zu korrigieren ist.

Zu (b): Die Frage, mit welchem Recht unterschiedlichen Strukturaspekten, wie sie durch bestimmte CJ-Gruppen repräsentiert sind, der gleiche Stufenwert zugewiesen werden kann, läßt sich positiv nur dann beantworten, wenn und insoweit nachgewiesen werden kann, daß die für bestimmte Stufen konstitutiven soziomoralischen Perspektiven sich als sinnstiftende Einheiten konsistent und trennscharf durchhalten.

Die Frage nach der hinreichenden Konsistenz und Trennschärfe der im Kohlberg-Manual vorgenommenen Stufenbestimmungen läßt sich auf normativ-analytischer und begriffs- bzw. bedeutungsanalytischer Ebene abklären. Diese Klärung bedarf unter dem Gesichtspunkt einer eventuellen inhaltlichen Verengung bestimmter Stufen- bzw. Strukturbestimmungen freilich der Ergänzung durch Untersuchungen auf empirisch-analytischer Ebene.

Da Kohlbergs standardisiertes Auswertungsinstrument (Manual) die Aufgabe hat, die Subsumption bestimmter Interview-Urteile unter bestimmte Stufenstrukturen im voraus eindeutig zu regeln und festzulegen, ist mit dem Aufzeigen bestimmter Problembereiche auf CJ-Ebene,

d.h. mit dem Nachweis problematischer Vereindeutigungen bzw. Bedeutungsunterstellungen sowie Bedeutungsverschiebungen nicht nur die Validität des Meßinstruments, sondern zugleich auch die Konstruktvalidität des Kohlbergschen Stufenmodells in Frage gestellt. Um mit der Validität des Meßinstruments nicht zugleich auch die Konstruktvalidität des Kohlbergschen Stufenmodells in Frage stellen zu müssen, bedarf es einer "Entkoppelung" der Validierungsebenen, die nur gelingen kann, wenn anstelle der präjudikativen Funktion eines standardisierten atomistischen Auswertungsverfahrens eine lediglich heuristische Funktion des Auswertungsmanuals tritt, d.h. wenn hinsichtlich der genannten Problemzonen der erforderlichen hermeneutischen Kontextualisierung der Bedeutung Rechnung getragen wird.

Auf dem Niveau konventioneller Moral sind folgende Problemzonen der Kohlbergschen Manualkonstruktion nachweisbar:

Auf Stufe 4 kommt es innerhalb eines Dilemmas zu Bedeutungsunterstellungen. Konkrete Rollenerwartungen werden - in problematischer Vereindeutigung und Festlegung im voraus - im Sinne einer Systemperspektive interpretiert, indem unterstellt wird, daß sie als "gesellschaftlich delegierte" bzw. "gesellschaftlich definierte" zu verstehen sind.

Im (personalen) Joe-Dilemma, Stufe 4, Issue: "Versprechen", können folgende CJs dies exemplarisch illustrieren:

CJ#26:

"[Joe should refuse to give his father the money] because the father is not asking for the money for the sake of the family welfare; or because the father does not want the money for something vital to the household.

[It is important to the father to keep his promise] because otherwise he will contribute to the breakdown of the family law ...; or because dictatorial actions will cause disharmony, animosity, or irresponsibility in the family."

"Stage structure: This judgment reflects a conception of the family as a social institution within which there exist informal moral and legal codes, authorities responsible for the welfare of the group, and a kind of social order analogous to that which exists on the broader societal level ..."

CJ#27:

"[Joe should refuse to give his father the money] because the father is misusing or abusing his authority, or is not using his power responsibly; or because the father isn't worthy of a son's normal obligation of respect."

"Stage structure: The focus in this judgment is on the father's breaking of his promise to his son as an indication that the father is not acting as a responsible and conscientious parent. Although there is an acceptance of the legitimacy of the father's societally defined authority role, this acceptance is conditional upon the father's acting as a person of integrity, so that he deserves the son's respect for his authority, ... the legitimacy of Stage 4 fixed social authority is balanced by a concern for the father's fulfilment of his responsibility."

CJ#28:

"[Joe should refuse to give his father the money] because Joe is now reaching maturity and should not be dominated by his father; or because Joe is reaching the age where he should be on an equal footing with his father."

"Stage structure: In this judgment there is a recognition of a societally defined role of the parent as an authority over a minor which is balanced against a consideration that as the son approaches adulthood a norm of equity ought to operate between them."

Im Dilemmatavergleich (personal vs. transpersonal) kommt es auf den Stufen 4 und 3 zu dilemma- und issueabhängigen Bedeutungsverschiebungen: Dies kann in exemplarischer Weise durch den Vergleich des Korea-Dilemmas mit dem Joe-Dilemma illustriert werden:

Im transpersonalen Korea-Dilemma treten auf Stufe 4 bei insgesamt 11 situationsspezifischen CJs 6 CJs auf, die eine "konkrete Systemperspektive" in dem Sinne repräsentieren, daß auf die mit bestimmten sozialen Positionen verknüpften (normativen) Erwartungen im Hinblick auf ein übergeordnetes soziales Ganzes Bezug genommen wird (s. Issue "Lebensqualität", CJ#27, #28, #29, #31, #32; Issue "Lebenserhaltung" CJ#22); insgesamt 5 CJs repräsentieren eine "abstrakte Systemperspektive", die dadurch gekennzeichnet ist, daß von konkreten Rollenerwartungen abstrahiert und nur den allgemeinen Funktionsbedingungen sozialer Institutionen und Systeme Rechnung getragen wird (s. Issue "Lebensqualität", CJ#26, Issue "Lebenserhaltung", CJ#18, #19, #20, #21). - Issuespezifisch zeigt sich eine deutliche Abhängigkeit von den Issues (autonom vs. heteronom): während von insgesamt 6 CJs bei dem autonomen Issue "Lebensqualität" 5 CJs eine konkrete Systemperspektive und nur ein CJ eine abstrakte Systemperspektive repräsentieren, tritt bei insgesamt 5 CJs im heteronomen Issue "Lebenserhaltung" bei 4 CJs eine abstrakte Systemperspektive und nur bei einem CJ eine konkrete Systemperspektive auf.

Im Joe-Dilemma treten bei insgesamt 12 situationsspezifischen CJs insgesamt 6 CJs auf, die eine konkrete Systemperspektive repräsentieren (s. Issue "Versprechen", CJ#24, #26; Issue "Autorität", CJ#20, #21, #22, #23), und nur ein CJ, das durch eine abstrakte Systemperspektive (Issue "Versprechen" CJ#23) gekennzeichnet ist.

Kritisch ist festzuhalten, daß insgesamt 5 CJs weder eine abstrakte noch eine konkrete Systemperspektive, sondern eine Beziehungsperspektive repräsentieren (s. Issue "Versprechen", CJ#25, #27, #28, #31, #33). Auch beim (personalen) Joe-Dilemma zeigen sich klare Issue-Effekte: Von insgesamt 8 CJs bei dem Issue "Versprechen" repräsentieren 5 (!) CJs überhaupt keine Systemperspektive, sondern eine Beziehungsperspektive. Die Trennschärfe, mit der im Joe-Dilemma, Issue "Versprechen" , zwischen Stufe 3 und Stufe 4 differenziert wird, ist folglich mangelhaft. Neben den genannten 5 CJs (Non-

Systemperspektiven) sind beim Issue "Versprechen" 2 CJs durch eine konkrete Systemperspektive und ein CJ durch eine abstrakte Systemperspektive gekennzeichnet. Im Issue "Autorität" repräsentieren alle 4 CJs eine konkrete Systemperspektive.

Auf Stufe 3 tritt im Korea-Dilemma bei insgesamt 15 situationsspezifischen CJs in 5 CJs bereits eine - für Stufe 4 kennzeichnende - konkrete Systemperspektive auf (s. Issue "Lebensqualität": kein CJ; Issue "Lebenserhaltung": CJ#9, CJ#10, CJ#12, CJ#13, CJ#15); ein CJ (s. Issue "Lebensqualität", CJ#13) repräsentiert eine prosoziale Einstellung universalistischer Art und ist deshalb - im Rahmen einer erforderlichen Kontextualisierung der Bedeutung - eventuell im Sinne des Übergangs von konventioneller zu postkonventioneller Moral (Stufe 4/5) zu verstehen. Es zeigen sich innerhalb des (transpersonalen) Korea-Dilemmas issueabhängige Bedeutungsverschiebungen folgender Art: im Issue "Lebensqualität" sind von insgesamt 8 CJs 5 CJs durch prosoziale Einstellungen gekennzeichnet, 2 CJs repräsentieren Rollenerwartungen, die nicht in eine konkrete Systemperspektive integriert sind, und ein CJ ist durch eine universalistische Sichtweise charakterisiert. Im Issue "Lebenserhaltung" repräsentieren von insgesamt 7 CJs bereits 5 (!) CJs, die Stufe 3 zugeordnet worden sind, eine konkrete Systemperspektive; 2 CJs sind durch prosoziale Einstellungen partikularistischer Art gekennzeichnet.

Im (personalen) Joe-Dilemma sind auf Stufe 3 alle 14 CJs, die insgesamt auftreten, durch prosoziale Einstellungen oder durch normative Erwartungen gekennzeichnet, die an partikulare, enge Gefühlsbeziehungen gebunden bleiben. Rollenerwartungen, die in eine konkrete Systemperspektive integriert wären, treten nicht auf. Im Gegensatz zum (transpersonalen) Korea-Dilemma treten innerhalb des Joe-Dilemmas auf Stufe 3 keine issueabhängigen Bedeutungsverschiebungen auf, die die Stufenkriterien der Konsistenz und Trennschärfe als problematisch, d.h. als nicht hinreichend erfüllt erscheinen ließen.

Während sich auf konventionellem Niveau zwar die theoretischen, nicht aber die operationalen Stufenbestimmungen auf CJ-Ebene als unproblematisch erweisen, sind auf präkonventionellem Niveau ähnlich wie bei der Abgrenzung von konventionellem und postkonventionellem Niveau bestimmte CJ-Gruppen sowohl auf theoretischer als auch auf operationaler Ebene als problematisch zu kennzeichnen.

Auf den Kohlbergstufen 0, 1 und 2 ist die Bestimmung der Stufenstruktur inhaltlich verengt: auf Stufe 0 im Sinne der Interessen von Ego, auf Stufe 1 im Sinne der Erwartungen von Autoritätspersonen und auf Stufe 2 im egoistisch-instrumentalistischen Sinne. Wir gehen in Übereinstimmung mit Keller davon aus, daß Kohlberg bei dieser Definition der präkonventio-

nellen Stufen inhaltliche Verengungen vornimmt, die einerseits empirisch nicht haltbar und andererseits unter kognitiv-strukturtheoretischen Gesichtspunkten nicht notwendig sind (Keller, Eckensberger & von Rosen, 1989).

Wie problematisch Kohlbergs inhaltliche Verengungen auf präkonventionellem Niveau sind, manifestiert sich in den Untersuchungen von Damon, 1984; Keller, Eckensberger & v. Rosen 1989; Eisenberg, 1979; Hoffman, 1984; Keller, 1989; Keller, 1990; Nunner-Winkler & Sodian, 1988; Turiel, 1983). Alle diese Untersuchungen belegen, daß schon kleine Kinder nicht nur an Interessen von Ego und Erwartungen von Autoritätspersonen, sondern auch empathisch an Interessen von Alter orientiert sowie regelorientiert in der Weise sein können, daß auch das Handeln bzw. moralisch nicht berechnete Regelsetzungen von Autoritätspersonen an der Gültigkeit von elementaren, allgemeinen Regeln ihren Maßstab haben; außerdem können sie ihre Entscheidungs- bzw. Normenbegründungen in gewisser Weise auch schon in beziehungs- und gerechtigkeitsorientierter Form entwickeln.

Unter strukturtheoretischen Gesichtspunkten treten - auf den Punkt gebracht - vor allem folgende Problemzonen auf:

Stufe 0 ist in Kohlbergs spätem, als gültig veröffentlichten Stufenmodell nicht mehr vertreten. Es bleibt deshalb die Frage offen, wie sich die prämorale Struktur bestimmt. Wir schlagen folgende allgemeine, d.h. bereichsübergreifende Strukturbestimmung für Stufe 0 vor: Zwischen dem, was jeweils faktisch erwartet bzw. faktisch geboten, verboten oder erlaubt wird und normativen Erwartungen, an denen das eigene Handeln wie das Handeln anderer (auch das von Autoritätspersonen) gemessen werden muß, kann noch nicht differenziert werden. Was "gut" ist, wird im Sinne der faktischen Präferenzen von Ego oder Alter verstanden.

Stufe 1: Das Konzept der Egozentrität, durch das für Kohlberg Stufe 1 definiert ist, bleibt unklar. Es bleibt offen, ob damit gemeint ist, daß (a) selbst auf der Ebene der Konfliktwahrnehmung noch nicht zwischen den unterschiedlichen Standpunkten (Interessen und Gefühlslagen) von Ego und Alter differenziert werden kann, oder ob (b) zwar bei der Konfliktwahrnehmung bereits zwischen unterschiedlichen Interessenlagen differenziert werden kann, bei der Konfliktlösung aber nur die eigenen Interessen Berücksichtigung finden. Unklar bleibt auch, wie Personen mit dem Konflikt zwischen den (faktischen) Erwartungen und Regelsetzungen von Autoritätspersonen und elementaren, allgemeinen Normen umgehen. Kohlberg läßt zudem außer Acht, daß die Verbindlichkeit moralischer Normen nicht selbstevident ist, sondern schon von kleinen Kindern unter empathischen Gesichtspunkten begründet werden kann (vgl. Turiel, 1983).



Wir schlagen folgende allgemeine Strukturbestimmung von Stufe 1 vor: Auf Stufe 1 wird im Gegensatz zu Stufe 0 zwischen faktischen und berechtigten Erwartungen differenziert. Nicht nur das eigene Handeln, sondern auch das Handeln von Autoritätspersonen muß sich an der "verabsolutierten" Gültigkeit von elementaren, allgemeinen Normen messen lassen. Moralische Normen, die mit elementaren Regeln identifiziert werden ("Du sollst nicht lügen, nicht stehlen, nicht betrügen, nicht töten usf.) gelten unverbrüchlich. Ihre Gültigkeit kann nicht hinsichtlich unterschiedlicher Anwendungsbedingungen eingeschränkt oder außer Kraft gesetzt werden.

Stufe 2: Die Konfundierung von Struktur- und Inhaltsaspekten beruht darauf, daß Kohlberg die Reversibilität der Perspektivenübernahme nicht hinreichend formal, sondern in instrumentalistisch-egoistisch verengter Weise faßt.

Wir schlagen folgende Bestimmung der universellen Tiefenstruktur von Stufe 2 vor:

Im Unterschied zu Stufe 1 läßt Stufe 2 eine differenzierte, psychologisch relativ sensible Berücksichtigung von unterschiedlichen Problem-, Interessen- und Gefühlslagen zu, und zwar auch dann, wenn diese im Widerspruch zu elementaren allgemeinen Normen steht. Die Relativierung elementarer Normen erfolgt unter dem Gesichtspunkt, daß Ego von Alter berechtigterweise erwarten kann, daß er sich in seine Lage versetzt und umgekehrt, allerdings noch nicht in der Form, daß die wechselseitige Berücksichtigung der unterschiedlichen Interessen - und Gefühlslagen im Sinne der Goldenen Regel selber zur Norm, d.h. zu einer Regel zweiter Ordnung gemacht würde (Stufe 3). Die Verbindlichkeit elementarer allgemeiner Normen kann im Hinblick auf unterschiedliche Anwendungsbedingungen in situationsspezifischer Weise modifiziert, d.h. eingeschränkt oder punktuell außer Kraft gesetzt werden. Dabei kann nicht nur auf isolierte, einzelne Interessenstandpunkte Bezug genommen werden, sondern unterschiedliche Teilnehmerperspektiven werden unter Reziprozitätsgesichtspunkten (im Sinne negativer oder positiver Vergeltung) und kompromißorientierter Ausbalancierung der konfligierenden Standpunkte koordinierbar. Die Koordination von unterschiedlichen Beteiligtenperspektiven, wie sie auf Stufe 2 hinsichtlich der Fähigkeit zu wechselseitiger Rollenübernahme vorgenommen werden kann, erfolgt noch nicht im Rekurs auf normative Erwartungen bzw. auf konkrete soziale Rollen, wie sie für die Aufrechterhaltung "guter", besonderer Beziehungen konstitutiv sind, in denen bestimmte Personen zueinander stehen. Die Koordination von Beteiligtenperspektiven geschieht nicht im Lichte einer überparteilichen Beobachterperspektive, die Individuen als Mitglieder einer sozialen Gemeinschaft miteinander teilen (Stufe 3).

Auf präkonventionellem Niveau, vor allem auf Stufe 2, sind hinsichtlich der operationalen Stufenbestimmungen Bedeutungsverschiebungen und Bedeutungsunterstellungen nachweisbar, die exemplarisch in Bezugnahme auf das Heinz-Dilemma und das Euthanasie-Dilemma illustriert werden sollen.

Bei Kohlberg sind auf Stufe 2 sowohl in Übereinstimmung mit seiner inhaltlich verengten theoretischen Stufenbestimmung CJs angesiedelt worden, die instrumentelle Entscheidungsbegründungen (s. Heinz, Issue: "Leben", CJ#6; Issue: "Gesetz", CJ#8, Euthanasie, Issue: "Lebensqualität," CJ#6; Issue: "Lebenserhaltung", CJ#7) enthalten als auch solche CJs, die, im Widerspruch zur theoretischen Stufenbestimmung Kohlbergs, empathische Entscheidungsbegründungen beinhalten (s. z.B. Euthanasie-Dilemma, Issue: "Lebensqualität", CJ#3, CJ#4). Außerdem sind einige CJs weder klar empathisch noch instrumentalistisch verstehbar (Heinz, Issue: "Leben", CJ#3, CJ#5; Euthanasie, Issue: "Lebensqualität", CJ#2).

Von "empathischer Entscheidungsbegründung" ist sinnvollerweise dann die Rede, wenn der Proband sich mitfühlend in die Interessen-, Gefühls- und Problemlage der von einer Handlungsentscheidung betroffenen Personen versetzt und primär deren Ansprüchen Rechnung zu tragen sucht. Der Betroffene kann dabei der vor eine (moralische) Entscheidung gestellte Akteur oder eine andere Person sein. Während im Falle des autonomen Issues im Euthanasiedilemma (s. Issue: "Lebensqualität", CJ#3, CJ#4, CJ#5) wie auch im Heinzdilemma (s. Issue: "Leben", CJ#3, CJ#7) relativ häufig empathische Entscheidungsbegründungen auftreten, kommen beim heteronomen Issue (Heinz, Issue: "Gesetz", Euthanasie, Issue: "Lebenserhaltung") in diesen Dilemmata ausschließlich nicht-empathische Entscheidungsbegründungen auf.

Empathische Entscheidungsbegründungen sowie Entscheidungsbegründungen, deren Interpretation offen bleibt, werden bei Kohlberg im instrumentalistischen Sinne umgedeutet.

Euthanasiedilemma, Issue: "Lebensqualität", CJ#3: "[The doctor should give the woman the drug] if he wants to help her get rid of her pain; or because she wants to die, doesn't want to be in pain anymore, or wants to get rid of her pain."

Stage Structure: "This judgment focusses on the woman's hedonistic or pragmatic desires or needs as a sufficient justification for the mercykilling ... the woman's desire to be rid of pain is a sufficient justification and is not stated in terms of the empathy or concern of a moral person considering her desire."

Nicht nur die Strukturbestimmungen, sondern auch die Inhaltsbestimmungen Kohlbergs erweisen sich als problematisch. Wie schon erwähnt, erhebt Kohlberg den Anspruch einer strikten analytischen Trennung der Struktur- und Inhaltsaspekte moralischen Urteilens und ist nur um eine Rekonstruktion von Inklusionsbeziehungen bei CJs bemüht, die durch gleiche Normen und Elemente gekennzeichnet sind.

Unsere Manualanalyse führte zu dem Ergebnis, daß ganze Elementegruppen auf präkonventionellem Niveau noch nicht und auf postkonventionellem Niveau nicht mehr auftreten; lediglich auf konventionellem Niveau sind alle 4 Elementegruppen vertreten: Auf präkonventionellem Niveau kommen die beiden Gruppen "perfektionistische Orientierung an Idealen des Selbst und des sozialen Zusammenlebens" sowie "Fairneßorientierung" nicht vor, mit Ausnahme des Elements der "Ausbalancierung von Perspektiven". Auf postkonventionellem Niveau tritt weder die Elementegruppe "Aufrechterhaltung der normativen Ordnung" noch die Gruppe "utilitaristische Folgeorientierung" in Erscheinung, mit Ausnahme des Elements "Folgen für die soziale Gruppe."

Damit ist aber die Vergleichbarkeit der unterschiedlichen Stufen moralischer Urteilsfähigkeit und eine über mehrere Stufen sich erstreckende Rekonstruktion von Inklusionen unter inhaltlichen Aspekten von vornherein erheblich eingeschränkt: Inklusionen können bei allen 4 Elementegruppen, von den erwähnten Ausnahmen abgesehen, maximal über 2 Niveaus auftreten. Bei den ersten beiden Elementegruppen ("Aufrechterhaltung der normativen Ordnung" und "utilitaristische Folgeorientierung") umfassen die möglichen Inklusionen das präkonventionelle und konventionelle Niveau bzw. die Stufen 1 bis 4, bei den Elementegruppen "perfektionistische Orientierung" und "Fairneß-Orientierung" schließen die möglichen Inklusionen nur das konventionelle und postkonventionelle Niveau bzw. die Stufen 3 bis 5 ein.

Die maximal möglichen Inklusionen werden bei den meisten Elementen und der durch diese Elemente gekennzeichneten CJs nur in sehr eingeschränkter Form ausgeschöpft, mit Ausnahme weniger Elemente bzw. Norm-Elementekombinationen, z.B. dem Element "dient der menschlichen Würde und Autonomie", das in Kombination mit der Norm "Leben" über 5 halbe Stufen (3, 3/4, 4 und 5) auftritt.

Um die beiden Defizite - in starkem Maße stufenabhängige Elemente und Inklusionen - eliminieren zu können, schlagen wir folgende Inhaltsdimensionen vor:

1. personorientiert: (a) egoorientiert, (b) empathisch-alterorientiert;

2. beziehungs- bzw. gruppenorientiert: definiert sich hinsichtlich der Aufrechterhaltung bestimmter ("guter") Beziehungen (a) zu Gleichgestellten, (b) zu Autoritätspersonen;
3. gehorsams- und autoritätsorientiert: definiert sich hinsichtlich der Erwartungen, Gebote, Verbote etc. von Autoritätspersonen;
4. regel- bzw. normorientiert;
5. gerechtigkeitsorientiert.

### 2.1.3 Weiterentwicklung des Präferenztests von Rest

Folgende Fragen, die sich zunächst in sachbezogener Form, d.h. in Bezug auf die o.g. Problemzonen, zugleich aber auch in methodischer Weise hinsichtlich ihrer Entscheidungsfähigkeit bzw. "Testbarkeit" stellen, müssen in empirischen Untersuchungen geklärt werden:

Inwieweit kommt es zwischen den (präjudikativen) Standards rationaler Rekonstruktion, wie sie in den mit den CJs verknüpften "Stufenstrukturen" (stage structures) und "Inklusionen" zum Ausdruck kommen, und den bedeutungskonstituierenden Sichtweisen bzw. Standards der Subjekte selbst zu Übereinstimmungen und Diskrepanzen? Die Frage macht, um in hinreichend differenzierter, empirisch überprüfbarer Form entscheidbar zu werden, zwei methodische Schritte erforderlich, die zu einer Erweiterung des Kohlbergschen Stufenmodells führen:

Erster Schritt: In einem Verfahren, das der Prüfmethode von Rest (1983) entlehnt ist, werden die Probanden, die zu einem Meßzeitpunkt  $t_x$  auf einer Kohlbergstufe (n) als der höchsten ihnen erreichbaren (Kompetenz-)Stufe oder aber als der modalen Stufe stehen, unter normativ-analytischen Gesichtspunkten mit problematischen CJs konfrontiert, die unterschiedliche Stufen (z.B.  $n-1$ ,  $n$ ,  $n+1$ ,  $n+2$ ) repräsentieren. Zwischen diesen Stufen haben die Probanden zu wählen, und zwar zunächst einmal ohne eine moralische Begründung der Präferenzen vorzunehmen.

Mit diesem ersten Schritt läßt sich folgende Frage verknüpfen: Inwieweit stimmen die faktischen Stufenpräferenzen der Subjekte selbst mit der Kohlbergschen Stufenfolge überein? Darüber hinaus kann sinnvollerweise untersucht werden, inwiefern sich die auf der Ebene faktischer Präferenzen feststellbaren Übereinstimmungen bzw. Diskrepanzen als dilemma- oder frageabhängig, stufenabhängig oder personabhängig erweisen.

Zweiter Schritt: Der zweite Schritt führt über Rests Erfassung faktischer Präferenzen hinaus, da es um die Erfassung moralischer Begründungen für bestimmte Vorrangordnungen geht.

Mit dem zweiten Schritt lassen sich folgende Fragen verknüpfen: Inwiefern stimmen die moralischen Gesichtspunkte, mit denen die Subjekte selbst begründen, warum eine Stufe  $x$  einer Stufe  $y$  vorzuziehen oder nicht vorzuziehen ist, mit den (präjudikativen) Standards rationaler Rekonstruktion überein? Inwiefern sind die "Inklusionen", die in Kohlbergs Auswertungsmanual in problematischer Vereindeutigung als Inklusionen präjudiziert sind, als von den Subjekten selbst hergestellte Stufendifferenzen zu verstehen (z.B. die Inklusion beim Element "dient der menschlichen Würde und Autonomie" im Heinz-Dilemma)?

Die mit dem zweiten Schritt verknüpfbaren Fragestellungen können in folgender Weise allgemein formuliert werden: Bedeuten Argumente einer Kohlbergstufe ( $n$ ) in den Augen der Subjekte das gleiche wie Argumente der Stufe ( $n-1$ ) bzw. ( $n+1$ ), oder bedeuten sie etwas anderes?

Auch auf der Ebene der von den Subjekten selbst hergestellten normativen Präferenzrelationen kann sinnvollerweise untersucht werden, ob bzw. inwiefern sie sich als kontextabhängig, als stufenabhängig oder als personabhängig erweisen.

## 2.2 Das Problem der Vermittlung von hermeneutischer und psychometrischer Forschungsstrategie

Im Rahmen der für Kohlbergs Forschungsansatz konstitutiven Vermittlungsproblematik von rationaler und phänomenologischer Rekonstruktion einerseits, von hermeneutischer und psychometrischer Forschungsstrategie andererseits, ist nicht nur eine einseitige Betonung rationaler Rekonstruktion, sondern auch die einseitige Betonung einer psychometrischen Auswertungsstrategie zu kritisieren.

Kohlbergs Rekonstruktion der Entwicklung moralischer Urteilsfähigkeit will an unterschiedlichen Maßstäben gemessen sein: an den Standards der Moralpsychologie als hermeneutisch gewendeter empirischer Sozialwissenschaft, zugleich aber an psychometrischen Standards im Sinne der empirisch verfahrenen "mainstream"-Psychologie (s. Kohlberg, 1984). - Damit stellt sich das bei Kohlberg nicht hinreichend erörterte und abgeklärte Problem, ob bzw. inwiefern die unterschiedlichen Standards von Hermeneutik und psychometrischer Verfahrensweise miteinander zu vereinbaren sind.

Kohlberg geht davon aus, daß der psychometrische Zugriff auf die erhobenen Daten ein dekontextualisiertes, atomistisches Auswertungsverfahren zur Voraussetzung hat, das isolierte, einzelne Urteile des Interviews unter isolierte, einzelne CJs und die dadurch

repräsentierten Stufenwerte subsumiert - unter hermeneutischen Gesichtspunkten ein gewiß nicht unproblematisches Unternehmen.

Die Problematik von Kohlbergs vorrangig psychometrisch orientierter Forschungsstrategie zeigt sich in hermeneutischer Hinsicht unter vier Gesichtspunkten: (a) bei der Manualkonstruktion, (b) bei den Regeln zur Ermittlung von CJ-, Issue- und Formwerten, (c) bei der Analyse von ausgewerteten Daten und (d) bei der Interpretation der analysierten Daten.

Zu (a): Wie problematisch Kohlbergs atomistisches Auswertungsverfahren ist, wurde im unvermeidbaren Vorgriff auf Ergebnisse unserer Manualanalyse bereits hervorgehoben, ohne dabei jedoch den hermeneutischen Gesichtspunkt einer Kontextualisierung der Bedeutung einzelner Aussagen näher zu erörtern.

Bei Kohlbergs Manualkonstruktion kommt es zu Bedeutungsunterstellungen bzw. zu vereinheitlichenden Bedeutungszuschreibungen, durch die im voraus (präjudikativ) festgelegt wird, wie bestimmte Interviewurteile zu verstehen sind, welche Stufenwerte sie jeweils repräsentieren. Von der Notwendigkeit einer Kontextualisierung der Bedeutung wird dabei abgesehen.

Im Rahmen eines dekontextualisierten Auswertungsverfahrens werden die Funktionen nicht erfaßt, die einzelne Argumentationsschritte in relativ komplexen Argumentationsgängen haben.

Unter dem (hermeneutischen) Gesichtspunkt problematischer Vereindeutigung bzw. Bedeutungsunterstellung kann auf folgende CJ-Gruppen, deren Bedeutung sich nur im Argumentationszusammenhang erschließt, hingewiesen werden: Problematisch sind vor allem Entscheidungsbegründungen unter Bezugnahme auf konkrete Rollenerwartungen, bei denen es offen bleibt, ob sie Bestandteile einer Systemperspektive (Stufe 4) sind. Problematisch bzw. nicht im voraus klar abgrenzbar sind auch der Stufe 4 zugeordnete CJs, bei denen Konzepte der Integrität des moralischen Selbst (Achtung, Ehre, Gewissen) begründungsrelevant werden. Problematisch sind ferner gewisse CJs, die Wertehierarchisierungen beinhalten, da eine begründete Zuordnung zu Stufe 3 oder zu Stufe 4 eine nur im hermeneutischen Kontext mögliche Klärung der Frage voraussetzt, inwiefern jeweils im einzelnen Fall zwischen der sprachlichen Elaboration und dem kognitiven Entwicklungsniveau unterschieden werden kann. Es versteht sich, daß nur in den Fällen, in denen vorherrschend Argumentationsmuster durch die moralische Sichtweise von Stufe 3 oder 4 zu kennzeichnen sind oder wo zumindest die schwächere, freilich selbst problematische Bedingung erfüllt ist, daß eine

Systemperspektive, Wertehierarchisierungen oder ein Integritätskonzept im Sinne von Stufe 4 vorliegt, eine Vereindeutigung problematischer Äußerungen bzw. der entsprechenden CJs im Sinne der Assimilation an Stufe 4 vorgenommen werden kann.

Zu (b): Zu einer minimalen Ausschöpfung des Interviewmaterials kommt es nicht nur aufgrund der schmalen empirischen Konstruktionsbasis des Manuals, sondern auch aufgrund der Regeln für die Kodierung von CJs. Diese Regeln sehen vor, daß der Kodierprozeß prinzipiell abgeschlossen ist und keine Rate-Scores (guess scores) zu ermitteln sind, wenn bei einem bestimmten Issue ein CJ aufgefunden worden ist und weitere CJs nicht auffindbar sind. Durch diese einseitige Ausrichtung an (im Manual vorgesehenen) CJs wird häufig unter strukturellen Gesichtspunkten bedeutsames Material verschenkt.

Kohlbergs Manualkonstruktion verfolgt das Ziel, eine eindeutige Zuordnung von Interviewurteilen zu bestimmten Criterion Judgments zu ermöglichen, folglich subjektive Interpretationsspielräume von vornherein einzuengen und zu minimieren. Sie erzeugt aber gerade dadurch in der Anwendung des Manuals als Komplementärphänomen relativ große Interpretationsspielräume in Form von "guess-scores", da eine rigorose Einengung der Interpretationsspielräume zu einer minimalen Ausbeute an einstuftbaren Intervieweinheiten geführt hätte. Guess-Scores lassen die Auswertung von mehrdeutigem Material zu und von Interviewmaterial, das im nicht-behandelten Issue anfällt. Guess-Scores kommen dann ins Spiel, wenn bei einem bestimmten Issue kein Interviewmaterial auffindbar ist, das einem bestimmten CJ entspricht. Dies kann sowohl einer fehlerhaften Interviewführung als auch den Eigenheiten und Idiosynkrasien individueller Argumentationen geschuldet sein (s. Colby & Kohlberg, 1987, S. 181ff.). Wie wenig eine psychometrisch standardisierte Interviewauswertung greift, zeigt sich daran, daß Guess-Scores ca. 40% der vergebenen Issue-Werte ausmachen.

Problematisch ist auch das Gewichtungssystem, auf dem die Verrechnung von Issue-Werten zu Form-Werten beruht. Guess-Scores und die durch CJs repräsentierten Stufenwerte werden im Hinblick darauf, ob ein Issue mit der Handlungspräferenz (action choice) des Probanden übereinstimmt, im Rahmen eines Punktesystems unterschiedlich bewertet. Stufenwerte, die auf der Grundlage von Guess-Scores vergeben werden, erhalten einen Punkt, während die durch CJs repräsentierten Stufenwerte im Falle des gewählten Issue 3 Punkte und im Falle des nicht-gewählten Issue 2 Punkte erhalten. Dieses Punktesystem spiegelt die Überzeugung, daß ein Issue, das die Handlungspräferenz des Probanden wiedergibt, kompetenznahe Leistungen eher repräsentiert als das alternative, nicht-gewählte Issue - eine Annahme, die einer empirischen Überprüfung jedoch nicht standhält. De Vries & Walker (1986) konnten nachweisen, daß die Stufenwerte zum Teil beim nicht-gewählten Issue höher liegen.

Zu (c): Nicht nur Kohlbergs dekontextualisiertes Auswertungsverfahren, sondern auch die Auswertung des Interviewmaterials auf hohem Aggregationsniveau ist unter hermeneutischem Gesichtspunkt problematisch.

Von Kohlberg sind bei der Auswertung des Interviewmaterials wie schon erwähnt drei verschiedene Aggregationsniveaus der Datenauswertung vorgesehen, nämlich CJ-Werte, Issue-Werte und Form-Werte.

Die atomistische Auflösung relativ komplexer Argumentationszusammenhänge in isolierte, einzelne Interviewurteile und deren Zuordnung zu entsprechenden CJs ist bereits kritisiert worden. Trotz der von uns geltend gemachten kritischen Vorbehalte gegenüber Kohlbergs Auswertungsverfahren ist positiv hervorzuheben, daß die CJs für die Sicht- und Argumentationsweise der Subjekte selbst noch relativ sensibel sind. Bedauerlicherweise bleibt die CJ-Ebene bei der Analyse der ausgewerteten Daten jedoch weitgehend unberücksichtigt (s. Kap. 3, "Ergebnisse auf empirisch-analytischer Ebene"). Dies liegt möglicherweise daran, daß bei Kohlberg auf CJ-Ebene eine hinreichende Interraterreliabilität nicht hergestellt werden konnte - ein Mangel, der Kohlbergs atomistisches Auswertungsverfahren auch hinsichtlich psychometrischer Standards problematisch erscheinen läßt.

Im Gegensatz zu den CJ-Werten sind Issue- und Formwerte (mehr oder weniger) künstliche Werte in dem Sinne, daß sie für die moralische Sicht- und Argumentationsweise der Subjekte nicht sensibel sind. Issue-Werte ergeben sich aus der Aggregation von CJ-Werten; die Issue-Werte werden ihrerseits - anhand der 6 Issue-Werte in einer Form - zu Formwerten aggregiert. Die erforderliche Interraterreliabilität wurde bei Kohlberg lediglich auf dem Aggregationsniveau der Formwerte überprüft.

Zu (d): Kohlbergs Analyse der ausgewerteten Daten seiner Längsschnittstudie reduziert sich auf die Feststellung, daß die Anteile von Stufen formabhängig sind. Die Abhängigkeit der Stufenwerte von bestimmten Dilemmata bzw. Dilemmatypen (z.B. personal vs. transpersonal) und von Issues wird von Kohlberg nicht untersucht. Genau dieser Untersuchungsgesichtspunkt ist aber hinsichtlich der uns interessierenden Frage, ob und inwiefern sich das moralische Urteilen als situationsadäquate Urteilsbildung verstehen läßt, von entscheidender Bedeutung.



Wie greifen bei der Erfassung bzw. Rekonstruktion situationsadäquater Urteilsbildung hermeneutische und psychometrische Forschungsstrategien ineinander?

Zunächst kann auf empirisch-psychometrischer Ebene untersucht werden, ob sich die Stufenanteile - und zwar vor allem hinsichtlich der konventionellen und postkonventionellen Stufen - als abhängig von bestimmten Dilemmatypen (personal vs. transpersonal) erweisen. Es ist plausibel, dabei von der Annahme auszugehen, daß personale Dilemmata stärker mit Stufe 3 und transpersonale Dilemmata stärker mit Stufe 4 verknüpft sind. Dies schließt logischerweise die Annahme ein, daß im Falle eines gemeinsamen Auftretens von konventionellen und postkonventionellen Argumentationsstufen postkonventionelle Stufen bei personalen Dilemmata signifikant häufiger mit Stufe 3 und bei transpersonalen Dilemmata mit Stufe 4 verknüpft sind. Die Annahme hingegen, daß sich die beiden Dilemmatypen auch hinsichtlich der Häufigkeit postkonventioneller Stufen unterscheiden, hat keine größere Plausibilität für sich als die entgegengesetzte Annahme.

Diese Annahmen lassen sich dadurch spezifizieren, daß sie unter differentiellen (typologischen) Gesichtspunkten interpretiert werden: Dabei gehen wir davon aus, daß sich die Unterschiede als personenabhängig erweisen. Einer Personengruppe (a), für die ein relativ flexibler, situationsadäquater Gebrauch der im Zuge der Entwicklung zur Verfügung stehenden Argumentationsstufen charakteristisch ist, wird eine Personengruppe (b) gegenüber gestellt, bei der dies nicht der Fall ist.

Um innerhalb des Systems von Kohlberg diese Annahmen in subjektsensibler Form überprüfen zu können, ist eine Datenauswertung auf CJ-Ebene erforderlich. Es wurden deshalb anhand des (nicht revidierten) Kohlbergmanuals umfangreiche Rekodierungen vorgenommen, die allerdings noch der Ergänzung bedürfen. Darüber hinaus kann sinnvollerweise im Rahmen einer kohlbergnahen und begrenzten Manualrevision auf Stufe 4 zwischen konkreter und abstrakter Systemperspektive differenziert werden.

Das empirisch Feststellbare wird aus kognitiv-internen Gründen verstehbar erst, wenn bestimmte Befunde (z.B. kontext- und personenabhängiger Stufenvariabilität) unter einem genuin moralpsychologischen normativen Konzept, z.B. unter dem von uns geltend gemachten Gesichtspunkt situationsadäquater Urteilsbildung, auf die Sicht- und Argumentationsweise der Subjekte selbst rückbezogen werden können, wobei in typologischer Hinsicht zwischen den genannten Personengruppen zu unterscheiden ist.

Personengruppe (a) ist dadurch gekennzeichnet, daß von den Kohlbergstufen 3, 3/4, 4, 4/5 und 5 flexibel und situationsadäquat Gebrauch gemacht wird. In diesem Fall kommt es also nicht zu einer schematischen (undifferenzierten) Assimilation der moralischen Entscheidungsbegründung an eine konsolidierte, stark ausgeprägte Modalstufe (n) im Sinne generalisierender Anwendung und Stabilisierung. Eine Ausnahme wäre gegeben, wenn auf Stufe 4 von der notwendigen Differenzierung zwischen konkreter und abstrakter Systemperspektive abstrahiert wird - eine problematische Voraussetzung, die bereits kritisiert wurde.

Personengruppe (b) hingegen ist durch eine relativ unflexible, schematische Verwendung von verfügbaren Argumentationstufen bzw. -formen charakterisiert. Die moralischen Entscheidungsbegründungen werden mit hoher interner Konsistenz an eine konsolidierte, stark ausgeprägte Modalstufe (n) im Sinne einer generalisierenden Anwendung und Stabilisierung assimiliert.

Zum psychometrischen Instrumentarium, das die Erfassung intraindividuelle Stufenvariabilität ermöglicht, zunächst einmal nur soviel: Zusätzlich zum Modus als Parameter für die zentrale Tendenz im Rahmen der intraindividuellen Stufenvariabilität können in Anlehnung an Walker und Taylor als Streuungsmaße die Parameter "positive" und "negative Mischung" und "Gesamtmischung" sowie "positiver" und "negativer Überhang" herangezogen werden (vgl. Walker & Taylor, 1991), und zwar unter Berücksichtigung nicht nur der reinen Stufen oder unter Verrechnung von Zwischenstufen zu reinen Stufenwerten, wie dies bei Walker und Taylor geschieht, sondern unter Berücksichtigung der Zwischenstufen als solchen. Die Anwendbarkeit dieses psychometrischen Instrumentariums auf die Kohlbergdaten bleibt allerdings begrenzt: die Stufen 1 und 5 treten als Modalstufen überhaupt nicht auf.

Im Projekt sind wir in die Rekonstruktion situationsadäquater Urteilsbildung bei solchen Fällen eingetreten, die auf der Ebene einzelner Issue-Werte postkonventionell werden. Auf CJ-Ebene stößt der Versuch einer Kohlbergimmanenten Rekonstruktion aufgrund der Datenlage von vornherein an eine zumindest vorläufig nicht aufhebbare Grenze: bei den uns zur Verfügung gestellten Auswertungsbögen Kohlbergs fehlen die meisten Fälle mit postkonventionell werdenden Formwerten; sie sind anscheinend verloren gegangen (!).

Auf dem Aggregationsniveau von Issue-Werten treten aufgrund der Regeln, mit denen CJ-Werte zu Issue-Werten verrechnet werden, postkonventionelle Argumentationsstufen nurmehr in reduzierter Form in Erscheinung (siehe hierzu das folgende Kapitel). In allen Fällen, die

auf dem Aggregationsniveau von Issue-Werten postkonventionell werden, ist ein hohes Ausmaß von Inkonsistenz bzw. Stufenstreuung festzustellen. Postkonventionelle Stufenwerte sind nämlich in jedem dieser Fälle mit verschiedenen konventionellen Stufenwerten (3, 3/4, 4) verknüpft. Durch Einzelfallanalysen kann untersucht werden, inwieweit die selbst auf dem Aggregationsniveau von Issue-Werten noch konstatierbaren weitreichenden Inkonsistenzen als (mehr oder weniger) komplexe Argumentationsmuster rekonstruierbar sind, die unter dem Gesichtspunkt situationsadäquater Urteilsbildung als kohärente Sinn Ganzheiten verstanden werden können. Eine moralische Argumentationsweise, die mit hoher Konsistenz im Sinne der Kohlbergstufe 5 an abstrakten Prinzipien und Verfahrensregeln oder auf konventionellem Niveau an einer abstrakten Systemperspektive (Stufe 4) orientiert wäre, könnte nur entwickelt werden auf Kosten einer situationsadäquaten Urteilsbildung (vgl. Edelstein & Noam 1982). In beiden Fällen könnte nämlich Konsistenz nur dann hergestellt werden, wenn von den spezifischen Konfliktsituationen (den Dilemmatypen und den für sie jeweils kennzeichnenden Handlungsbedingungen) abstrahiert würde, d.h. wenn bloß assimilatorisch, ohne Akkommodation an die Vielfalt besonderer Situationsumstände, ein kognitives Schema reproduziert würde.

Von situationsadäquater Urteilsbildung kann sinnvollerweise nur die Rede sein, wenn abstrakte Begründungsschemata vermieden bzw. wenn die o.g. Abstraktionen in gewisser Weise wieder rückgängig gemacht werden, und zwar durch die Verknüpfung mit konkreteren Sichtweisen, sei es im Sinne einer Beziehungsperspektive (Stufe 3), sei es im Sinne einer konkreten Systemperspektive (Stufe 4).

Situationsadäquate Urteilsbildung definiert sich im Hinblick auf spezifische Situations- und Handlungsbedingungen, an denen eine sinnvolle Anwendung moralischer Normen ihren Maßstab hat: sie berücksichtigt die spezifischen Rechte und Pflichten, wie sie für bestimmte Personen, die in bestimmten Beziehungen zueinander stehen, konstitutiv sind und sie berücksichtigt dabei die besonderen Individuallagen (Probleme, Interessen und Gefühle) der an einem Konflikt beteiligten Personen.

Situationsadäquate Urteilsbildung berücksichtigt aber auch, daß moralische Handlungsbedingungen, wie sie sich im Rahmen sozialer Institutionen und Systeme definieren, nicht angemessen in Analogie zu den Handlungsbedingungen in intimen, interpersonellen Beziehungen (d.h. im Lichte einer reinen Beziehungsperspektive) verstanden werden können, sondern nur im Lichte einer Systemperspektive. Inwiefern dabei nicht nur eine konkrete, in hierarchischer Integration organisierte Systemperspektive, sondern auch eine abstrakte Systemperspektive als situationsadäquat zu begreifen ist, hängt davon ab, inwieweit

restriktiven institutionellen sowie systemischen Funktions- und Handlungsbedingungen realitätsgerecht Rechnung getragen werden muß.

Eine hermeneutische Rekonstruktion von mehr oder weniger komplexen Argumentationsmustern hat den o.g. Kriterien situationsadäquater Urteilsbildung Rechnung zu tragen, gleichgültig, ob dies innerhalb eines gegebenen Dilemmas geschieht, oder ob unterschiedliche Dilemmata und bestimmte Argumentationsformen miteinander in Beziehung gesetzt werden. Dabei geht man davon aus, daß die Bedeutung, die einzelne Argumentationsschritte haben, sich in beiden Fällen durch die Funktion definieren, die bestimmte Argumentationsschritte in komplexeren Argumentationszusammenhängen haben.

Hinsichtlich der dilemmainternen Rekonstruktion ist eine Illustration wahrscheinlich hilfreich. Folgender Fall läßt sich konstruieren (Joe-Dilemma, Issue: "Vertrag"): Zunächst wird abstrakt im Sinne der prinzipienorientierten Moral von Stufe 5 auf universelle Grundrechte (der Freiheit und Gleichbehandlung) Bezug genommen, ohne sich auf die spezifischen Rechte und Pflichten, wie sie sich in einer Vater-Sohn-Beziehung definieren, einzulassen. Diese Abstraktion wird in einem zweiten Argumentationsschritt korrigiert, d.h. die spezifische Vater-Sohn-Beziehung findet bei der Entscheidungsbegründung die - im Sinne situationsadäquater Urteilsbildung - erforderliche Berücksichtigung, und zwar im Lichte einer Beziehungsperspektive (Stufe 3). Natürlich wäre eine umgekehrte Reihenfolge der Argumentationsschritte ebenfalls im Sinne situationsadäquater Urteilsbildung zu verstehen, obwohl in diesem Fall ihre Verknüpfung im Sinne rekursiver Begründung erfolgt:

Stufe 5: CJ#39

"[Joe should refuse to give his father the money, or property should be respected] because Joe is an individual whose personal rights are of equal value to his father's; or because each person as a free individual is entitled to enjoy the use of property as long as it does not infringe upon the like rights of others."

Stufe 3: CJ#13

"[The father should not ask for the money] because that breaks up the relationship or because he is causing Joe to lose faith in him or in the relationship."

### 2.3 Analyse des Standard Issue Scoring-Manuals

Wichtige Ergebnisse der Manualanalyse sind bei der Klärung der in den Kapiteln 2.1 und 2.2 behandelten Probleme bereits vorweggenommen worden.

(a) Hinsichtlich des Strukturaspekts moralischer Argumentationsstufen wurde auf die zwei Problembereiche der "Bedeutungsunterstellung" und der "Bedeutungsverschiebung" in Kohlbergs Manualkonstruktion hingewiesen. Außerdem wurde, um den Übergang von konventioneller zu postkonventioneller Moral bzw. die Abgrenzung zwischen diesen beiden Niveaus moralischer Urteilsbildung in der notwendigen Differenziertheit und Trennschärfe bestimmen zu können, ein Klassifikationssystem eingeführt, das zum einen die unzureichende Repräsentation gewisser Begründungsstrukturen im Übergang von konventioneller zu postkonventioneller Moral in Kohlbergs Manual sichtbar macht und zum anderen gewisse Umstufungen nahelegt.

(b) Hinsichtlich der Inhaltsaspekte und Struktur-Inhaltsrelationen, die für bestimmte Kohlbergstufen kennzeichnend sind, ist bereits in Abschnitt 2.1 darauf hingewiesen worden, daß nur auf dem konventionellen Niveau alle 4 Elementegruppen auftreten. Auf dem präkonventionellen Niveau treten nur die ersten beiden Elementegruppen ("Aufrechterhaltung der normativen Ordnung" und "utilitaristische Folgenorientierung") und auf postkonventionellem Niveau nur die beiden letzten Elementegruppen ("perfektionistische Orientierung" und "Fairneßorientierung") auf. Diese Stufenabhängigkeit der Inhaltsdimensionen hat aber, wie bereits in 2.1 hervorgehoben wurde, zur Folge, daß auch Kohlbergs Anspruch, die Stufenstrukturen entwicklungslogisch im Sinne "hierarchischer Integration" bestimmen und aufeinander beziehen zu können, auf der Ebene der operationalen Stufenbestimmungen im Manual erheblich eingeschränkt werden mußte. Die Spannweite der Inklusionen umfaßt maximal 5 halbe Stufen.

In den Abschnitten 2.1 und 2.2 wurden lediglich das Verhältnis der CJs zu den Stufenstrukturen (stage structures) und zu den Inklusionen berücksichtigt. Außerdem wurde nur auf die reinen Stufen Bezug genommen.

In diesem Kapitel sollen auch die prototypischen Beispiele (match examples) und kritischen Indikatoren (critical indicators) im Verhältnis zu den CJs, die kritischen Indikatoren im Verhältnis zu den Stufenstrukturen sowie die Zwischenstufen in die Betrachtung einbezogen werden.

Zunächst ist festzuhalten, daß die CJs ungleich verteilt sind, und zwar in Abhängigkeit von

- (a) den Niveaus und Stufen
- (b) den Issues (autonom vs. heteronom)
- (c) den Fragetypen (konzeptuell vs. situationsspezifisch).

Zu (a): Es zeigt sich, daß bei allen Dilemmata die Stufen 1, 1/2, 4/5 und 5 unterrepräsentiert sind im Verhältnis zu Stufe 2 auf präkonventionellem Niveau und zu den Stufen 2/3, 3 und 4 auf konventionellem Niveau.

Zu (b): Auf präkonventionellem Niveau treten bei den Stufen 1 und 1/2 beim heteronomen Issue wesentlich mehr situationsspezifische Entscheidungsbegründungen auf als beim autonomen Issue, während auf postkonventionellem Niveau bei den Stufen 4/5 und 5 beim autonomen Issue deutlich mehr situationsspezifische Entscheidungsbegründungen vorhanden sind als beim heteronomen Issue. Zum Beispiel treten im Heinz-Dilemma beim Issue "Leben" auf den Stufen 1 und 1/2 nur zwei situationsspezifische CJs auf, beim Issue "Gesetz" hingegen kommen 5 situationsspezifische CJs vor. Auf den Stufen 4/5 und 5 gibt es beim Issue "Leben" 6 CJs und beim Issue "Gesetz" nur 1 CJ. Beim heteronomen Issue ("Gesetz") findet sich im Heinz-Dilemma kein CJ auf Stufe 5. Auch bei den heteronomen Issues im Joe-, Judy-, Bob-Karl- und Euthanasie-Dilemma sind auf Stufe 5 keine situationsspezifischen CJs vorhanden.

Zu (c): Während bei allen personalen Dilemmata bei den Issues nicht nur situationsspezifische, sondern auch konzeptuelle Frage hinreichend repräsentiert sind, treten beim Heinz-Dilemma, Issue "Leben", und beim Euthanasie-Dilemma, Issue "Lebensqualität" sowie im Korea-Dilemma, Issue "Lebensqualität" keine oder nur wenige CJs auf, die konzeptuelle Fragen repräsentieren würden, obwohl in all diesen Dilemmata zu den genannten Issues konzeptuelle Fragen im Katalog der Standardfragen aufgeführt sind. Beim Officer-Brown-Dilemma, Issue "Gewissen" und beim Korea-Dilemma, Issue "Lebenserhaltung" sind konzeptuelle Fragen überhaupt nicht gestellt worden und treten deshalb auch im Manual nicht auf.

Wenn die Annahme, daß bei konzeptuellen Fragen signifikant höhere Stufenwerte auftreten als bei situationsspezifischen Fragen, sich als richtig erweisen sollte, so hätte dies folgende Konsequenz: bei allen Issues, die keine konzeptuellen Fragen enthalten, käme es von vornherein zu einer Benachteiligung, d.h. zu einer relativen Absenkung der erwartbaren Stufenwerte.

Die Ergebnisse hinsichtlich des Verhältnisses (a) von prototypischen Beispielen (match examples) sowie kritischen Indikatoren und CJs und (b) von kritischen Indikatoren und Stufenstrukturen (stage structures) lassen sich wie folgt auf den Punkt bringen: Sowohl die Passung von prototypischen Beispielen und CJs als auch die Übereinstimmung von CJs und kritischen Indikatoren erweisen sich, von vereinzelt Ausnahmen abgesehen, als unproblematisch, nicht hingegen das Verhältnis von kritischen Indikatoren und Stufenstrukturen.

Inwieweit die Stufenstrukturen die ihnen zugeordnete Funktion struktureller Evaluation (s. Colby & Kohlberg, 1987) in Kohlbergs Auswertungsverfahren erfüllen und im Vergleich zu den kritischen Indikatoren präzisere und strukturtheoretisch ausgewiesene Stufenzuordnungen ermöglichen, ist fragwürdig. Die Stufenstrukturen stellen nämlich häufig nur wörtliche Wiederholungen oder abstraktere Fassungen der bereits bei den kritischen Indikatoren erfolgten Kennzeichnungen dar. Sie gehen insofern über die "Oberflächenaspekte", die durch die kritischen Indikatoren erfaßt werden sollen, im Grund nicht hinaus. Dies kann ein Beispiel aus dem Heinz-Dilemma, Issue "Leben", Stufe 3, illustrieren. CJ#15:

Criterion Judgment: "[Heinz should steal the drug, even if he doesn't love his wife] out of gratitude or appreciation; OR because she has shared her life with him, and the least he can do is to save her.

Stage Structure: "In this judgment obligation is based on a relationship of mutual sharing. The sharedness of experience creates not only feelings of affection, but a generalized obligation."

Critical Indicators: "Required for a match is an appeal to either: (a) the prior sharing; or (b) the gratitude Heinz should express."

Sofern die Stufenstrukturen im Vergleich zu den kritischen Indikatoren neue, zusätzliche Bestimmungen enthalten, erweisen sich diese häufig als theoretische Überinterpretationen bzw. als Bedeutungsunterstellungen, z.B. im folgenden Fall: Heinz-Dilemma, Issue "Leben", Stufe 4, CJ#23.

Criterion Judgment: "[Heinz should steal the drug] because his wife can contribute to society."

Stage Structure: "In this judgment there is a stage 4 systems' view in which society is an object and has value independently of the mass of people making it up. An individual's life gains value through the functional contribution the individual can make the social system."

Critical Indicators: "Required for a match is an appeal to the saved person's ability to: (a) contribute to society, or (b) help in the advancement or betterment of society."

Als problematisch erweisen sich in Kohlbergs Manualkonstruktion auch die Strukturbestimmungen der Zwischenstufen. Die Zwischenstufen machen einen hohen Anteil der CJs aus. Ihr theoretischer Status bleibt aber unklar.

Die Zwischenstufen lassen sich in 3 Gruppen klassifizieren:

(a): Nur wenige CJs, die Zwischenstufen repräsentieren, sind positiv als qualitativ eigenständige Stufen bestimmt: z.B. wie CJ#22, im Heinz-Dilemma, Issue "Leben", Stufe 3/4.

Criterion Judgment: "[Heinz should steal the drug] because if he really loves her, then it would be right from his viewpoint, or then she would be important enough to him to risk going to jail; OR if he feels strongly enough that saving her life is more important than obeying the law."

Stage Structure: "The concern is with Heinz's subjective feelings in terms of which the value of one set of social expectations (to his wife) may outweigh another (to law, society). What is right, especially for Heinz, is determined by what Heinz feels is right, especially in the case of strong feelings".

(b) Relativ häufig werden die Zwischenstufen als Begründungsstrukturen bestimmt, die nicht mehr für eine Stufe (n) und noch nicht für eine Stufe (n+1) kennzeichnend sind, wobei nicht klar ist, inwiefern sie noch gewisse Bedeutungskomponenten der Stufe (n) sowie auch bereits gewisse Bedeutungskomponenten der Stufe (n+1) enthalten: z.B. im Heinz-Dilemma, Issue "Gesetz", Stufe 4/5, CJ#28.

Criterion Judgment: "[It is important to obey the law] because law makes possible the regularity or consistency which yields freedom and security for all individuals."

Stage Structures: "This judgment goes beyond the stage 4 conception of legal consistency and lawfulness as necessary for social order and maintenance of the social system toward an appreciation of the importance of such social goods as freedom and security for individual



members of the society. However, there is not yet a clear expression of the stage 5 concept that these benefits represent fundamental human rights".

(c) Am problematischsten ist die Berücksichtigung von Zwischenstufen, die als Ergebnis mangelhafter Interviewführung interpretiert oder auf Schwierigkeiten bei der Einstufung zurückgeführt werden, z.B. beim Heinz-Dilemma, Issue: "Leben", Stufe 2/3, CJ#8.

Criterion Judgment: [Heinz should steal the drug for a wife or a friend] if he is desperate; OR because he wouldn't have much choice; OR because his wife (friend) means a lot to him."

Stage Structure: "This justification is based on either a concern that Heinz is not to be blamed for stealing since he had no antisocial motive (Stage 3), or a consideration that Heinz's stealing would be an instrumentally reasonable thing to do (Stage 2). In this justification, the Stage 3 vs. Stage 2 notion of approval cannot be distinguished. If probed, the justification is usually identifiable as either Stage 2 (CJs#4,7) or included within Stage 3 points (CJs#12, 16).

#### 2.4 Das mehrdimensionale Modell

Kohlbergs Stufenmodell der Entwicklung moralischer Kompetenz bleibt in mehrfacher Weise begrenzt: zum einen schon dadurch, daß vorrangig die Entwicklung moralischer Urteilsfähigkeit und nicht gleichgewichtig die Entwicklung moralischer Handlungsfähigkeit zum Thema gemacht und erfaßt worden ist. Wie die Entwicklung moralischer Urteilsbildung im Kontext moralischen Handelns erfolgt, wie die Bedeutung, die sie in diesem Kontext besitzt bzw. die Rückwirkung, die das Handeln auf die moralische Entscheidungsfindung und deren Begründung hat, unter moralpsychologischen Gesichtspunkten zu verstehen ist, bleibt eine offene Frage. Moralische Urteile, die integrativer Bestandteil realen moralischen Handelns sind, sind in ihrer Struktur und Funktion nicht schlicht gleichzusetzen mit moralischen Urteilen, die sich auf hypothetische Dilemmata beziehen; und moralisches Handeln ist angemessen nicht als die bloße - mehr oder weniger konsistente oder inkonsistente - Umsetzung von vorab entwickelten Zielsetzungen und deren moralischen Begründungen zu verstehen. Für alle Konfliktparteien gleichermaßen akzeptable bzw. legitime Handlungsintentionen, Ziele, Mittel und Folgen definieren sich im Kontext realen moralischen Handelns angemessen nur im Rahmen einer Handlungskoordination, die (a) von den Subjekten als Zentren eigener selbstbestimmter Urteils- und Willensbildung vorgenommen wird, und zwar (b) in primär verständigungsorientierter Einstellung. Die gemeinsame Konstruktion von entsprechenden

Formen der Handlungskoordination wird bei Kohlberg aber z.B. beim Heinzdilemma von vornherein unterbunden.

Im Kontext des Kohlbergschen Ansatzes wiegt indessen vielleicht schwerer, daß noch auf der Ebene des moralischen Urteilens selbst eine Verengung vorgenommen wird: Kohlbergs Stufenmodell ist rein präskriptiv-deontisch orientiert. Aufgrund der von Kohlberg vorgenommenen, problematischen Gleichsetzung moralischer Urteile, die eine rational rekonstruierbare Struktur aufweisen, mit Gerechtigkeitsurteilen bleibt sein Stufenmodell ein beschränktes, eindimensionales Modell. In diesem Stufenmodell der Entwicklung moralischer Urteilsfähigkeit bleibt unbegriffen oder wird zumindest nicht hinreichend berücksichtigt, (a) daß nicht nur deontische Urteile, sondern auch Verantwortlichkeitsurteile (vgl. Blasi, 1983; Ropohl, 1994; Keller, 1994) eine rationale Struktur besitzen, d.h. in (mehr oder weniger) intersubjektiv verbindlicher Form begründbar sind und deshalb die Differenzierungen und Beziehungen, die im Zuge der Entwicklung der moralischen Urteils- und Handlungsfähigkeit zwischen diesen beiden Urteilsformen hergestellt werden, ein möglicher Gegenstand rationaler Rekonstruktion sind. (b) Dies gilt auch für Formen der Differenzierung und Beziehung, wie sie zwischen präskriptiven und deskriptiven soziomoralischen Urteilen im Verlaufe der Entwicklung der moralischen Urteils- und Handlungsfähigkeit von den Subjekten selbst hergestellt werden (vgl. Keller, 1994; Keller & Edelstein, 1986; Nunner-Winkler & Sodian, 1988). (c) Schließlich wird auch der rationalen Struktur der moralischen Handlungskoordination, die ihren Maßstab an der Fähigkeit und Bereitschaft zu verständigungsorientiertem Handeln hat, nicht in adäquater, d.h. handlungstheoretischer Form Rechnung getragen (vgl. Selman, 1986; Eckensberger & Reinshagen, 1980; Habermas, 1983; Keller & Reuss, 1984; Keller & Reuss, 1986).

Selbst auf der Ebene deontischer Urteile erfolgt eine weitere Verengung: zwischen Geltungs- und Anwendungsfragen wird nicht in hinreichend systematischer Form differenziert.

Diese kritischen Gesichtspunkte führen über die bereits entwickelte Kritik an Kohlbergs Niveau- und Stufenbestimmungen moralischer Urteilsfähigkeit auf präskriptiv-deontischer Ebene hinaus und bringen neue, von Kohlberg nicht hinreichend berücksichtigte Dimensionen ins Spiel, Dimensionen, die für eine differenziertere Erfassung moralischer Urteils- und Handlungsfähigkeit gleichermaßen konstitutiv sind.

Das hier vorgeschlagene mehrdimensionale Modell moralischer Urteils- und Handlungsfähigkeit erhebt den theoretisch begründeten Anspruch, im Vergleich zu Kohlbergs eindimensionalem Modell moralischer Urteilsfähigkeit bestimmte "progressive Problemverschiebungen" im

Bereich kognitiv-strukturtheoretischer Forschungsentwicklung besser systematisch erfassen zu können. So können nicht nur präskriptive, sondern auch deskriptive soziomoralische Urteile untersucht werden, und sog. "real-life"-Dilemmata (die freilich entgegen dem terminologischen Anschein nicht in den Kontext realen moralischen Handelns eingebettet sind) sowie die Entwicklung des moralischen Selbst und die damit implizierte Konsistenz zwischen moralischem Urteilen und Handeln können thematisiert werden.

Auf empirisch-analytischer Ebene kann dieses mehrdimensionale Modell seine volle Problemlösungskapazität erst entfalten, wenn es sich auf darauf eigens zugeschnittene Instrumente der Datengewinnung (strukturelle Interviews) und der Datenauswertung (revidiertes und erweitertes Auswertungsmanual) stützen kann. Diese Instrumente müssen so beschaffen sein, daß sie es erlauben, Mehrfachkodierungen zu vermeiden und die unterschiedlichen Dimensionen als voneinander unabhängige zu behandeln, so daß dimensionsspezifische Stufenentwicklungen erfaßbar werden.

Unter differentiellem Gesichtspunkt stellt sich die Frage, ob bzw. inwiefern ein mehrdimensionales, handlungstheoretisches Modell der Entwicklung der moralischen Urteils- und/oder Handlungsfähigkeit eine differenziertere Typenrekonstruktion ermöglicht, als dies bei Kohlbergs dichotomischer Konstruktion von reinen Urteilstypen (autonom vs. heteronom) der Fall ist.

Das mehrdimensionale Modell erweitert Kohlbergs rein präskriptiv-deontisch orientiertes Entwicklungsmodell moralischer Urteilsfähigkeit, indem es in einem handlungstheoretischen Rahmen zwischen folgenden Dimensionen differenziert und in jeder Dimension polare, typenspezifische Merkmalsausprägungen unterscheidet:

#### 1. Deontische Urteile vs. Verantwortlichkeitsurteile:

- (a) *Übernahme der Verantwortlichkeit*
- (b) *Leugnung bzw. Verschiebung der Verantwortlichkeit*

#### 2. Deskriptive vs. präskriptive soziomoralische Urteile:

- (a) *Integriertes Motivationsverständnis:* (Kongruenz von Beweggründen und Rechtfertigungsgründen).
- (b) *Dissoziatives Motivationsverständnis:* (Beweggründe bzw. Motive, die erklären, wie jemand tatsächlich handelt und wie der Akteur sich dann fühlt und Rechtfertigungsgründe, die

begründen, warum eine Handlungsweise unter moralischen Gesichtspunkten als "richtig" zu bewerten ist, sind inkongruent).

### 3. Formen der Handlungskoordination:

(a) *verständigungsorientiert* (vorrangige Berücksichtigung übergeordneter gemeinsamer Interessen und Ansprüche)

(b) *strategisch orientiert* (einseitige Durchsetzung der Interessen und Ansprüche von Ego oder Berücksichtigung der Interessen und Ansprüche von Alter unter dem Gesichtspunkt der "Klugheit" - d.h. im wohlverstandenen Eigeninteresse eines Individuums.

### 4. Differenzierungsformen zwischen Geltungs- und Anwendungsfragen auf der Ebene deontischer Urteile:

(a) *situationsspezifische* Einschränkung, d.h. Modifikation oder Außerkraftsetzung deontischer Standards

(b) *schematische, nicht-situationsspezifische* Aufrechterhaltung deontischer Standards.

Es erfolgt jetzt eine Erläuterung dieser Dimensionen:

#### Zu 1: Deontische Urteile vs. Verantwortlichkeitsurteile

Hinsichtlich der Dimension der Verantwortlichkeitsurteile ist zumindest zwischen zwei polaren Merkmalsausprägungen zu differenzieren:

(a) Übernahme individueller oder kollektiver Verantwortlichkeit vs.

(b) Leugnung oder Verschiebung individueller oder kollektiver Verantwortlichkeit.

Kohlberg berücksichtigt in systematischer Form lediglich die deontische Dimension von Rechten und Pflichten: was richtig ist bzw. was in einer bestimmten moralischen Konfliktsituation getan werden soll, bestimmt sich nach Kriterien der Unparteilichkeit, der Universalisierbarkeit und der Zustimmungsfähigkeit.

Der Klasse von deontischen Urteilen, die durch die Erfüllung dieser Kriterien definiert sind, stellt Kohlberg in der Auseinandersetzung mit Gilligan (1984) eine Klasse von Urteilen gegenüber, in denen in partikularistisch beziehungsorientierter Weise eine Orientierung an Fürsorge und Verantwortlichkeit (care and responsibility), d.h. eine persönliche

Verantwortung für das Wohlergehen anderer, vor allem nahestehender Personen, zum Ausdruck gebracht wird. Kohlberg versteht Verantwortlichkeitsurteile als nicht universalisierbare Urteile. Daher kann er sie nicht an deontische Urteile rückbinden und danach fragen, inwiefern Verantwortung im Sinne einer individuellen oder kollektiven Form der tätigen Einlösung spezifischer Rechte und Pflichten verstanden wird. Es ist eine wichtige Frage, wie individuelle und kollektive Verantwortung einerseits, Eigen- und Fremdverantwortung andererseits im Laufe der Entwicklung definiert werden und ob und inwiefern eine typenspezifische Fokussierung auf individuelle oder gemeinsame Verantwortung sowie auf Eigen- oder Fremdverantwortung erfolgt.

So kann z. B. die Existenz eines Typus vermutet werden, der das Recht auf Leben universalistisch versteht, für die tätige Einlösung dieses Rechts aber individuell und partikularistisch nur diejenigen Personen verantwortlich macht, die in einer engen Beziehung zu der Person stehen, deren Leben es zu retten gilt.

Nicht thematisiert wird bei Kohlberg, daß Verantwortlichkeitsurteile auch Bestandteile von Entschuldigungs-, Rechtfertigungs- und Wiedergutmachungsstrategien sind (s. Keller, 1984). Bei dieser Klasse von Verantwortlichkeitsurteilen handelt es sich um "Urteile ex post", während Verantwortlichkeitsurteile, die sich auf die Zuständigkeit von Individuen oder Kollektiven für die tätige Einlösung bestimmter Rechte und Pflichten beziehen, "Urteile ex ante" darstellen. Bei Verantwortlichkeitsurteilen ex post geht es um Fragen der individuellen oder gemeinsamen kollektiven Zurechenbarkeit, um die Zuschreibung von Verantwortung für die Folgen einer Handlung oder deren Unterlassung, um Bewertungen des moralischen Selbst (von Individuen oder von Mitgliedern eines sozialen Kollektivs) hinsichtlich der Übernahme, der Leugnung oder der Verschiebung von Verantwortlichkeit im Falle von Normverletzungen. Entsprechende moralische Bewertungen können aus der Perspektive von "Ego" (als Akteur) oder aus der Perspektive von "Alter" als direkt oder indirekt Betroffenen erfolgen (s. Blasi, 1983; Döbert & Nunner-Winkler, 1978; Keller, 1984; Scott & Lyman, 1976; Sykes & Matza, 1957).

## Zu 2: Deskriptive vs. präskriptive soziomoralische Urteile

Kohlberg berücksichtigt bei der Interviewauswertung lediglich präskriptive Äußerungen. Im Gegensatz zu Kohlberg wird bei Keller & Edelstein (1986) sowohl bei der Entwicklung des Interviewleitfadens als auch bei der Interviewauswertung systematisch zwischen deskriptiven und präskriptiven (soziomoralischen) Urteilen differenziert. Tatsächlich getroffene Entschei-

dungen und deren Begründung ("wie würdest Du Dich entscheiden und warum?") werden zum Gegenstand "moralischer Bewertung" gemacht ("ist es richtig, sich so zu entscheiden?").

Wie differenzieren Personen selbst zwischen deskriptiven und präskriptiven soziomoralischen Urteilen und wie beziehen sie diese beide Motivklassen aufeinander? Im Hinblick auf die Rekonstruktion der Entwicklung des moralischen Selbst muß dabei zumindest zwischen den beiden folgenden Merkmalsausprägungen unterschieden werden:

(a) *Dissoziatives Motivationsverständnis*: Dabei handelt es sich um Beweggründe bzw. Motive, die erklären, wie jemand tatsächlich handelt und wie der Akteur sich dann fühlt und Rechtfertigungsgründe, die begründen, warum eine Handlungsweise (h) unter moralischen Gesichtspunkten als "richtig" zu bewerten ist und wie ein Akteur - im Falle von (h) und nicht-(h) als moralisches Subjekt sich fühlen sollte; Beweggründe und Rechtfertigungsgründe sind inkongruent (vgl. Keller & Edelstein 1986; Nunner-Winkler & Sodian, 1988).

(b) *Integriertes Motivationsverständnis*: Kongruenz von Motiven (Beweggründen) und Rechtfertigungsgründen.

### Zu 3: Formen der Handlungskoordination

Nicht nur das moralische Urteilen, sondern auch das moralische Handeln sowie die Beziehung zwischen moralischem Urteilen und Handeln müssen sich an bestimmten Rationalitätsstandards messen lassen. Diese Rationalitätsstandards definieren sich indessen in einem handlungstheoretischen Rahmen im Hinblick darauf, ob bzw. inwiefern sich die Akteure und die Betroffenen als Zentren eigener, selbstbestimmter Urteils- und Willensbildung verstanden und behandelt wissen sowie hinsichtlich der Fähigkeit und Bereitschaft zu verständigungsorientiertem Handeln.

Von zentraler Bedeutung für eine handlungstheoretische Rekonstruktion der Entwicklung moralischer Urteils- und Handlungsfähigkeit ist deshalb die Dimension der Handlungskoordination. Im Anschluß an Habermas unterscheiden wir analytisch zwischen (a) verständigungsorientierten und (b) strategischen Formen der Handlungskoordination. Eine Handlungskoordination ist unter moralischen Gesichtspunkten nur dann als "richtig" oder "angemessen" zu verstehen, wenn sie für alle Akteure bzw. Betroffenen gleichermaßen akzeptabel ist, d.h. wenn sie auf einer selbstbestimmten, gemeinsamen Konstruktion beruht.

Das Kohlbergsche Konzept der "idealen Rollenübernahme", d.h. einer vollständig reversibel und universalisierbar gewordenen Perspektivenübernahme zielt auf den gleichen "moralischen Gesichtspunkt" (moral point of view) als Rekonstruktionsprinzip moralischer Urteilsfähigkeit

wie das von Habermas geltend gemachte Konzept verständigungsorientierten Handelns, das seinen Maßstab am "Konsens rational motivierter Subjekte" hat (s. Habermas, 1983; 1991); Kohlbergs Konzept der Perspektivenübernahme bedarf allerdings, um die von ihm vorgenommene Verengung auf bloße Urteilsfähigkeit überwinden zu können, einer handlungstheoretischen Transformation bzw. der Übersetzung in eine "language of action". Verständigungsorientierte Formen der Handlungskoordination setzen voraus, daß die in einen Normen- oder Interessenkonflikt verwickelten Konfliktparteien gleichermaßen bereit und fähig sind, die für ein (moralisches) Dilemma jeweils konstitutiven, konfligierenden Handlungspläne angemessen zu rekonstruieren sowie Formen der Transformation und Koordination der einander ausschließenden Handlungspläne (Intentionen, Ziele, Mittel, Folgen) zu suchen, die im gemeinsamen Interesse aller Konfliktparteien liegen. Dies schließt die Erwartbarkeit einer wechselseitigen Transformation des moralischen Selbst mit ein. Die Zumutung einer einseitigen Transformation des moralischen Selbst von "Ego" oder von "Alter" hingegen kann als unberechtigt verstanden, d.h. sie kann kritisiert und zurückgewiesen werden.

Zu 4: Differenzierungsformen zwischen Geltungs- und Anwendungsfragen auf der Ebene deontischer Urteile:

In dieser Dimension muß zumindest zwischen zwei Differenzierungsformen unterschieden werden:

- (a) *situationsspezifische* Einschränkung, d.h. Modifikation oder Außerkraftsetzung deontischer Standards vs.
- (b) *schematische, situationsunspezifische* Aufrechterhaltung deontischer Standards.

Beide Differenzierungsformen zwischen Geltungs- und Anwendungsfragen werden im Rahmen einer Differenzierung und Koordination zwischen bestimmten Wert- und Geltungssphären entwickelt.

Obwohl der Standpunkt einer "autonomen, universalistischen Moral" eine Differenzierung zwischen den unterschiedlichen Wert- und Geltungssphären von Religion, Recht, Moral, Konvention usf. erforderlich macht, hat Kohlberg dieser Differenzierung nicht systematisch Rechnung getragen.

Im Rahmen einer Kohlbergrekonstruktion, der es primär um eine konstruktive Bewältigung der bei Kohlberg nachgewiesenen Problemzonen geht, ist deshalb zu untersuchen, ob und inwiefern von den Subjekten selbst eine Differenzierung und Koordination zwischen

bestimmten Wert- und Geltungssphären vorgenommen wird. Um welche Sphären - z.B. Religion und Moral, Recht und Moral, Konvention und Moral - handelt es sich dabei; welche Gewichtung bzw. Rangordnung wird zwischen den entsprechenden Sphären hergestellt - z.B. Recht rangiert vor Moral oder umgekehrt: Moralität kommt vor Legalität. Wie wird die Eigen-Sinnigkeit einer Sphäre jeweils verstanden? Wir erwarten, daß sich Individuen auch kulturspezifisch darin unterscheiden, welchen Wertsphären sie Vorrang einräumen.

Deontische Urteile untersucht Kohlberg vor allem unter dem Gesichtspunkt, wie bestimmte Rechte und Pflichten, die prima facie Gültigkeit beanspruchen, in bestimmten Situationen verstanden und angewandt werden. Er untersucht nicht, ob und inwiefern es Unterschiede des Entwicklungsniveaus gibt zwischen dem generellen Verständnis (Konzept) von Normen und der Begründung dafür, ob und in welcher Weise Normen in spezifischen Konfliktsituationen relevant sind.

Zu fragen wäre dabei vor allem, wie sich im Rahmen von personalen vs. transpersonalen Dilemmata die Fähigkeit entwickelt, zwischen Gesichtspunkten der Geltung und Gesichtspunkten einer situationsadäquaten Anwendung zu differenzieren. Es wäre darüber hinaus zu fragen, ob und inwiefern in typenspezifischer Weise bei bestimmten Personengruppen eine Fokussierung auf geltende Normen oder auf deren situationsadäquate Anwendung erfolgt, eine Differenzierung, die auch in Beziehung auf einen bestimmten (z.B. personalen) Dilemmatypus sinnvoll ist. Einer Personengruppe A, die der Auffassung ist, daß in einer Vater-Sohn-Beziehung ein vom Vater dem Sohn gegebenes Versprechen unbedingt (d.h. unter allen Umständen) gehalten werden muß, damit der Sohn nicht die Achtung oder das Vertrauen dem Vater gegenüber verliert, läßt sich z.B. eine Personengruppe B gegenüberstellen, die mit dem ausdrücklichen Hinweis auf besondere Situationsbedingungen Ausnahmen von der Regel macht bzw. gelten läßt - Ausnahmen im Interesse der Aufrechterhaltung einer guten Beziehung.

Eine systematische Berücksichtigung der für die Entwicklung moralischer Urteils- und Handlungsfähigkeit konstitutiven vier Dimensionen des multidimensionalen Modells macht eine Erfassung der für jede Dimension kennzeichnenden Merkmalsausprägungen erforderlich.

Individuelle Entwicklungsverläufe definieren sich in dieser Hinsicht durch bestimmte Profile, wobei ein Typus, der vorherrschend im Sinne der Merkmalsausprägung (a) gekennzeichnet ist, sich gegenüber einem vorherrschend durch (b) gekennzeichneten Typus durch die größere moralische Problemlösungskapazität auszeichnet.



Folgende Verlaufstypen lassen sich als "mögliche Sachverhalte" konstruieren:

Zunächst einmal ein Extremtyp (A), der während des gesamten Entwicklungsverlaufs konsistent vorherrschend (also zumindest in 3 Dimensionen) durch Merkmalsausprägungen im Sinne von (a) charakterisiert ist - im Gegensatz zu dem komplementären Extremtyp (B), bei dem konsistent vorherrschend Merkmalsausprägungen im Sinne von (b) auftreten; außerdem ein progressiver Typus (C), bei dem zu einem bestimmten Zeitpunkt der Entwicklung ein Wechsel von vorherrschend (b) zu vorherrschend (a) erfolgt und der dann in konsistenter Form vorherrschend (a) bleibt - im Gegensatz zu einem regressiven Typus (D), bei dem ein Wechsel von vorherrschend (a) zu vorherrschend (b) erfolgt und der dann auch konsistent vorherrschend (b) bleibt. Schließlich ist ein Typus (E) denkbar, bei dem (a) und (b) in unregelmäßiger Form alternieren, ohne daß ein konsistentes Verlaufsmuster progressiver oder regressiver Art auftreten würde.

### **3. Ergebnisse auf empirisch-analytischer Ebene**

#### 3.1. Kohlbergs einseitige Berücksichtigung hoch aggregierter Auswertungsdaten

Im folgenden setzen wir uns mit Kohlbergs Strategie bei der Analyse der Daten und mit seinen empirischen Annahmen zu den Stufen und Inhalten kritisch auseinander.

Kohlberg wertet die Argumentation der Probanden sowohl auf der Grundlage des "Standard Issue Scoring" (vgl. Colby, Kohlberg 1987, S. 151ff.) als auch auf der Grundlage des "Moral Type Coding Scheme" (vgl. ebd. S. 353ff) aus. Durch beide Verfahren werden Werte auf verschiedenen Aggregationsniveaus ermittelt. Im Rahmen des Standard Issue Scoring findet eine Aggregation der Stufenwerte von Criterion Judgments zu Issue-Werten und der Issue-Werte zu Form-Werten statt. Im Rahmen des Moral Type Coding Scheme werden die Werte von einzelnen Typenkriterien zu Dilemma-Werten und die Dilemma-Werte zu Form-Werten zusammengefaßt.

Von Kohlberg und von der an Kohlberg orientierten Forschung wurde das Standard Issue Scoring seit Ende der 70er Jahre in einer Vielzahl von Querschnitts- und Längsschnittstudien zum Einsatz gebracht, wobei auch verschiedene Studien außerhalb der USA durchgeführt wurden (vgl. Eckensberger 1995; Snarey, 1985; Snarey & Keljo, 1991). Neben Kohlbergs eigenen Längsschnittuntersuchungen in der Türkei und in Israel liegen Längsschnitt- und

Querschnittuntersuchungen aus acht weiteren Ländern vor (Bahamas, Deutschland, Finnland, Indien, Kenia, Neuguinea, Polen, Taiwan).

Nur in wenigen Studien wurde bislang das Verfahren zur Kodierung moralischer Typen verwendet. Allein in den drei Längsschnittstudien von Kohlberg sowie einer Untersuchung auf den Bahamas und in Taiwan ist seit Anfang der 80er Jahre die Argumentation der Probanden zusätzlich nach Typen kodiert worden (vgl. Logan, Snarey & Schrader, 1990).

Von den vorliegenden Studien, in denen das Standard Issue Scoring und das Verfahren zur Kodierung von Typen Anwendung fand, ist die USA-Studie am differenziertesten angelegt:

- Hier wird die moralische Urteilsbildung im Hinblick auf alle drei Interviewformen, also im Hinblick auf alle neun Dilemmata erhoben, während in den übrigen Studien meistens nur eine Interviewform (gewöhnlich Form A) benutzt wird.
- In der USA-Studie besteht die Untersuchungspopulation zwar nur aus männlichen Personen, doch wurde dem Einfluß einer großen Zahl von Variablen nachgegangen, z.B. dem Einfluß des sozioökonomischen Status der Eltern des Probanden, des sozioökonomischen Status des Probanden selbst, der Erziehungsstile der Eltern, der Ausbildung, der Religionszugehörigkeit, des soziometrischen Status, der Abwehr- und Bewältigungsstrategien des Probanden (vgl. Hart 1992). Sofern in den anderen Studien neben der Bedeutung des Faktors "Alter" auch die Bedeutung anderer Einflußgrößen erfaßt wird, beschränkt sich dies gewöhnlich auf die Variablen "Geschlecht des Probanden" und "sozioökonomischer Status der Eltern".
- In der USA-Studie wurde die Argumentation der Probanden längsschnittlich über 20 Jahre verfolgt, während in den übrigen Längsschnittuntersuchungen die intraindividuelle Veränderung der Urteilsbildung höchstens über einen Zeitraum von 12, meistens aber nur über einen Zeitraum von zwei oder drei Jahren erfaßt wird.

Aber auch in der differenziert angelegten USA-Studie zeigen sich wie in den anderen Studien gravierende Beschränkungen in der Analyse der auf der Basis des Standard Issue Scoring und des Typenscoring ermittelten Werte. Sowohl Kohlberg als auch die an Kohlberg orientierte Forschung haben sich vorwiegend auf die Analyse hoch aggregierter Auswertungsdaten konzentriert und die niedrig aggregierten Daten stark vernachlässigt:

- Die Analyse der Stufenentwicklung bestand weitgehend nur in einer Analyse der Form-Werte. Issue- und CJ-Werte wurden nur in wenigen Studien gezielt untersucht (vgl. de Vries & Walker 1986; Nisan & Kohlberg 1982; Walker & Taylor, 1991).
- Nur selten wurden bislang auf der Basis von Criterion Judgments die Elemente analysiert (vgl. Walker, 1989). Auch die dilemmaspezifischen Handlungspräferenzen wurden kaum zum Gegenstand der Forschung gemacht (vgl. Snarey, 1985). Die Normen wurden überhaupt noch nicht untersucht.
- Ausschließlich die Form-Werte wurden bei der Analyse der Typenentwicklung herangezogen.

So mangelt es an empirischen Befunden auf mehreren Ebenen:

- Es fehlen bislang Aussagen über die Stufenverteilung für die Issue- und CJ-Werte, Aussagen über das Ausmaß der intraindividuellen Stufenvariation sowie über Sequenzmuster bei den Issue- und CJ-Werten. Zudem fehlen Resultate zum möglichen Einfluß der einzelnen Dilemmata, Issues und Fragetypen sowie der Handlungspräferenzen auf die Urteilsstufen.
- Es mangelt an Befunden zu Stufenabhängigkeiten, intraindividuellen Konsistenzen, Sequenzmustern und Kontexteffekten in der Verwendung der Elemente und Normen und hinsichtlich der dilemmaspezifischen Handlungspräferenzen.
- Ebenso wenig existieren Ergebnisse über die Reliabilität des Standard Issue Scoring hinsichtlich der CJ-Werte und hinsichtlich der Kodierung von Normen und Elementen.
- Schließlich fehlen Informationen über die Typenentwicklung auf dem Aggregationsniveau der Dilemma-Werte und der Werte für Typenkriterien. Ebenso fehlen Resultate über die Reliabilität des Typenscoring für diese Werte.

Ein Ziel des Projekts war es, die vernachlässigte Mikroanalyse moralischen Urteilens voranzutreiben. Wir haben die Issue-Werte der USA-Stichprobe untersucht (3.2). Darüber hinaus legen wir erste Ergebnisse zur Interraterreliabilität der Kodierung von Criterion Judgments vor (3.3).

Noch nicht abgeschlossen ist eine Mikroanalyse der Entwicklung von Typen, die in Form einer Reanalyse von Auswertungsdaten der USA-Studie durchgeführt wurde. Die systematische Untersuchung der Normen, Elemente und Handlungspräferenzen steht ebenso aus wie die Analyse der Stufenwerte von CJs.

### 3.2 Ergebnisse zur Analyse der Issue-Werte

Es bestand zunächst die Absicht, für die Reanalyse der Stufenwerte der USA-Studie sowohl die Issue-Werte als auch die CJ-Werte heranzuziehen. Es zeigte sich aber bald, daß eine systematische Untersuchung der CJ-Werte auf der Grundlage der uns vorliegenden Auswertungsdaten nicht möglich ist. So fehlen für einige Fälle die Auswertungsbögen, in denen die Criterion Judgments verzeichnet sind. Darüber hinaus sind bei mehreren Fällen die Kodierungen nur schwer zu rekonstruieren. Wir haben uns deshalb auf die Untersuchung der Issue-Werte konzentriert.

Bei der Untersuchung der Issue-Werte standen drei Fragen im Vordergrund: die Frage nach dem Anteil postkonventioneller Stufen bei den Issue-Werten (3.2.1), das Problem der intraindividuellen Konsistenz der Issue-Werte (3.2.2) sowie die Frage nach der Kontextabhängigkeit der Stufenwerte (3.2.3).

#### *3.2.1 Der Anteil postkonventioneller Stufen bei den Issue-Werten*

Kohlberg und die an Kohlberg orientierte Forschung beanspruchen, mit Hilfe des moralischen Interviews und des Standard Issue Scoring das Urteilspotential, die "Kompetenz" der Person erfassen zu können (vgl. Colby & Kohlberg, 1987, S. 5). Die ermittelten Stufenwerte sind also als Kompetenzwerte zu verstehen.

Aussagen über die Kompetenz von Personen bzw. Stichproben liegen bislang allein auf der Ebene der Form-Werte vor. Die Aussagen beziehen sich hauptsächlich auf das Alter von 10 - 40 Jahren: es gibt nur wenige Untersuchungen zu Personen über 40 Jahren; die Urteilskompetenz unter 10 Jahren wurde überhaupt nicht erfaßt.

Betrachtet man die Ergebnisse der Studien, die seit Ende der 70er Jahre auf der Grundlage des Standard Issue Scoring durchgeführt wurden, fällt auf, daß sowohl Stufe 1 und 1/2 als auch Stufe 4/5 und 5 auf der Ebene der Form-Werte recht selten auftreten. Der geringe Anteil von Stufe 1 und 1/2 ist angesichts der erfaßten Altersspanne wenig überraschend. Überraschend dagegen ist der geringe Anteil der Postkonventionalität. So sind in den Studien zur Urteilsbildung in traditionellen Gesellschaften (z.B. Kenia, Bahamas) überhaupt keine postkonventionellen Stufen festgestellt worden. Selbst bei erwachsenen Personen, die in modernen westlichen Gesellschaften leben und in Städten aufgewachsen sind oder der Mittelschicht angehören, sind postkonventionelle Werte selten (vgl. Boyes & Walker, 1988; Eckensberger, 1995; Snarey, 1985; Snarey & Keljo, 1991).

In der USA-Studie wird nur einer einzigen Person (nämlich Fall 23) ein Form-Wert der Stufe 5 zugewiesen. Nur in 4 weiteren Untersuchungen wird bei Personen ein Kompetenzniveau der Stufe 5 ermittelt, wobei sich Stufe 5 nicht vor dem Alter von 30 Jahren findet: "Stage 5 ... proved to be quite rare. All subjects who scored as fully stage 5 were 30 years or older in age and came from only 3 countries. These included 2 socially elite subjects in India, 15 activists (war protesters or Kibbutz members) in Israel, and 10 subjects in the United States" (Snarey & Keljo, 1991, S. 404).

Stufe 4/5 wird in der USA-Studie nur von College-Absolventen erreicht und auch da nur von einem Teil der Absolventen (vgl. Colby & Kohlberg 1983; S. 56). In der USA-Studie und auch in den anderen bislang durchgeführten Untersuchungen findet sich Stufe 4/5 im allgemeinen erstmals ab dem Alter von 20 Jahren. Die Anteile von Stufe 4/5 übersteigen dabei nur in wenigen Studien 10% (vgl. Snarey & Keljo, 1991, S. 404f.).

Im Rahmen des frühen Sentence Scoring und Global Story Ratings hatte Kohlberg postkonventionelle Werte noch 10- und 13-jährigen zugewiesen. Bei den 16jährigen war Stufe 5 in der USA-Studie sogar die am häufigsten gebrauchte Stufe (vgl. Colby & Kohlberg, 1983; S. 68). Die Veränderung des Auswertungsverfahrens hat also zu einer massiven Ausdünnung postkonventioneller Werte geführt.

Mehrere Autoren weisen mittlerweile auf den geringen Anteil postkonventioneller Stufen kritisch hin. Sie leiten aus den erzielten Resultaten die Notwendigkeit ab, Kohlbergs Definition von Postkonventionalität zu modifizieren, sei es im Hinblick auf traditionale, nichtwestliche Gesellschaften (vgl. z.B. Snarey und Keljo, 1991) oder aber im Hinblick auf moderne Gesellschaften (vgl. z.B. Shweder, Mahapatra & Miller, 1987). Bei ihren kritischen Diagnosen stützen sich die Autoren dabei ausschließlich auf die Form-Werte. Es wird nicht in Betracht gezogen, daß der Anteil von Postkonventionalität auf der Ebene der Issue-Werte und der CJ-Werte möglicherweise um einiges höher liegt und der schwache Anteil bei den Form-Werten durch den Prozeß der Aggregation von Stufenwerten erklärt werden kann. Wir haben uns die Frage gestellt, wieweit in der USA-Studie das geringe Ausmaß postkonventioneller Form-Werte auf den Aggregationsprozeß zurückzuführen ist.

Durch die spezifische Form der Zusammenfassung von Stufenwerten im Rahmen des Standard Issue Scoring ist die Reduktion eines nicht unerheblichen Anteils postkonventioneller Werte theoretisch möglich.

Bei der Aggregation von CJ-Werten zu Issue-Werten und von Issue-Werten zu Form-Werten werden zunächst jeweils die Prozentanteile ermittelt, die die Stufe 1, 2, 3, 4 und 5 bei den CJs zu einem Issue bzw. bei den 6 Issues einer Form haben. Für die auf CJ- und Issue-Ebene auftretenden Zwischenstufen werden dabei keine eigenständigen prozentualen Häufigkeiten errechnet: die Anteile von Zwischenstufen werden jeweils als Anteile der reinen Stufen verrechnet. Sowohl bei der Zusammenfassung der CJ-Werte zu einem Issue-Wert als auch bei der Zusammenfassung von Issue-Werten zu einem Form-Wert werden nur reine Stufen berücksichtigt, die innerhalb eines Issues bzw. bei den 6 Issues einer Form einen Anteil von 25% oder mehr besitzen. Haben mehrere Stufen einen Anteil von 25% oder mehr, so ergeben sich jeweils Übergangsstufen.

Um einen Form-Wert der Stufe 5 zu erzielen, muß also die Person bei den meisten der 6 Issues Werte der Stufe 5 haben. Ist der prozentuale Anteil von anderen Stufen ähnlich hoch oder höher als der Anteil von Stufe 5, ergeben sich lediglich Übergangsstufen. Tritt ein Issue-Wert der Stufe 5 nur einmal auf und ist zugleich Stufe 5 auch nicht häufiger in Übergangsstufen enthalten, wird Stufe 5 aufgrund der 25%-Regel eliminiert. Denkbar sind darüber hinaus mehrere Konstellationen, bei denen Issue-Werte der Stufe 4/5 eliminiert werden.

Auch bei der Verrechnung von CJ-Werten zu einem Issue-Wert müssen starke Anteile von Stufe-5-Urteilen vorhanden sein, damit diese Stufe als reine Stufe und nicht als Übergangsstufe 4/5 in Erscheinung tritt. Weil zu einem Issue aber gewöhnlich nicht mehr als 4 CJs kodiert werden, dürften Werte der Stufe 5 jedoch selten vollständig neutralisiert werden, wogegen dies bei Werten der Stufe 4/5 häufiger möglich ist.

Bei der Bestimmung globaler Kompetenzwerte der Person durch Issue- und Form-Werte orientiert sich Kohlberg also nicht an Maximalwerten, sondern an Modalwerten. Eine als Maximalwert auftretende reine Stufe muß in konsolidierter Form vorliegen, um einen globalen Kompetenzwert, d.h. einen Issue- oder Form-Wert zu bilden. Die globalen Aussagen über die Kompetenz von Personen sind also bei Kohlberg am Kriterium der Konsolidierung von Strukturen orientiert.

Findet nun in der USA-Studie die theoretisch mögliche Unterschlagung von Postkonventionalität tatsächlich in stärkerem Ausmaß statt?

Eine Analyse der Stufenverteilungen für die in allen 3 Formen erzielten Issue-Werte (vgl. Teo, Becker & Edelstein, 1995; Tab 1) zeigt, daß die Gesamtanteile der Stufen 4/5 und 5 recht

gering bleiben. Ab dem Alter von 24 - 26 Jahren tritt Stufe 5 im Gegensatz zur Verteilung bei den Form-Werten bei einigen Personen zwar auf, aber die Anteile bei den einzelnen Altersgruppen bleiben unter 3 %. Stufe 4/5 findet sich erstmals ab dem Alter von 20 -22 Jahren in nennenswertem Umfang, kommt dabei aber über einen Anteil von 10% nicht hinaus. Damit unterscheiden sich die Prozentanteile von Stufe 4/5 und 5 auf der Ebene der Issue-Werte nur geringfügig von den entsprechenden Prozentanteilen auf der Ebene der Form-Werte (vgl. Colby & Kohlberg, 1983; S. 24, 46).

Der relativ geringe Gesamtanteil von Stufe 4/5 und 5 auf Issue-Ebene läßt den Schluß auf ein vermutlich seltenes Auftreten der Stufe 5 auf CJ-Ebene zu: Weil bei einer maximalen Anzahl von 4 CJs zu einem Issue CJ-Werte der Stufe 5 zumindest als Issue-Werte der Stufe 4/5 in Erscheinung treten müßten, der Häufigkeitsanteil aber selbst von Stufe 4/5 gering bleibt, kann vermutet werden, daß auch auf CJ-Ebene Werte der Stufe 5 nur selten auftreten. Eine systematische Analyse der Stufenanteile auf CJ-Ebene steht zwar noch aus, doch bestätigte eine erste Sichtung der vorliegenden CJ-Werte der Personen diese Vermutung. Zudem zeigte sich, daß auch Stufe 4/5 nicht sehr häufig vorhanden ist.

So ist also festzuhalten, daß der Gesamtanteil postkonventioneller Stufen auf der Ebene der Issue- und CJ-Werte bei weitem nicht so hoch ist wie es theoretisch möglich wäre.

Was die Stufen 4/5 und 5 auf der Ebene der Issue-Werte angeht, so ist zusätzlich festzuhalten, daß deren Anteile über die unterschiedlichen Issues und Dilemmata hinweg in etwa gleich niedrig sind, also spezifische Issues oder Dilemmata nicht für den relativ geringen Gesamtanteil verantwortlich sind.

Daß der Gesamtanteil postkonventioneller Stufen bei den Issue- und CJ-Werten recht gering ist, bedeutet aber nicht, daß keine postkonventionellen Maximalwerte unterschlagen werden. Betrachtet man die einzelnen Fälle, die postkonventionelle Issue-Werte erzielen, so zeigt sich, daß Maximalwerte der Stufe 4/5 und 5 bei mehreren Personen neutralisiert werden. So erhalten alle Personen mit einem oder zwei Maximalwerten der Stufe 4/5 nur konventionelle Form-Werte. Und allen Personen, die in einer Form einen oder zwei Maximalwerte der Stufe 5 erzielen, wird nur ein Form-Wert der Stufe 4/5 zugewiesen.

Beim 5. Meßzeitpunkt erzielen z.B. in Form A 3 Fälle, in Form B 3 Fälle und in Form C 6 Fälle einen oder zwei Maximalwerte der Stufe 4/5. Beim 6. Meßzeitpunkt erhalten in Form A 5 Fälle und in Form B 5 Fälle einen oder zwei Maximalwerte der Stufe 4/5 (Form C wurde

bei diesem Meßzeitpunkt nicht verwendet). Allen diese Fällen wird nur ein konventioneller Form-Wert der Stufe 3/4 oder 4 zugeschrieben.

Während Issue-Werte der Stufe 4/5 bereits ab dem 2. Meßzeitpunkt vorhanden sind, treten Issue-Werte der Stufe 5 erstmals zum 5. Meßzeitpunkt auf. Beim 5. Meßzeitpunkt ist in Form A und B jeweils ein Fall mit einem oder zwei Issue-Werten der Stufe 5 vorhanden; in Form C findet sich kein Fall mit Stufe 5. Beim 6. Meßzeitpunkt erreichen in Form A 3 Fälle und in Form B 4 Fälle einen oder zwei Issue-Werte der Stufe 5. Diese Fälle erhalten alle einen Form-Wert der Stufe 4/5.

Selbst bei einigen Fällen mit drei Maximalwerten der Stufe 4/5 bzw. 5 bilden die Maximalwerte nicht die Form-Werte. Zum Beispiel erzielt Fall 44 beim 4. Meßzeitpunkt dreimal Stufe 4/5 und zweimal Stufe 3 (ein Issue wurde nicht kodiert!), erhält aber nur einen Form-Wert der Stufe 3/4. Fall 2 erreicht beim 5. Meßzeitpunkt dreimal Stufe 5, zweimal Stufe 4/5 und einmal Stufe 4 und erhält einen Form-Wert der Stufe 4/5. Von insgesamt 6 Fällen mit drei Maximalwerten der Stufe 4/5 erzielen drei Fälle konventionelle Form-Werte. Von insgesamt 4 Fällen mit drei Maximalwerten der Stufe 5 werden drei Fällen Form-Werte der Stufe 4/5 zugewiesen. Dabei ist festzuhalten, daß Fälle mit mehr als drei Maximalwerten der Stufe 5 zu einer Form überhaupt nicht vorkommen. Fälle mit mehr als 3 Maximalwerten der Stufe 4/5 treten dagegen vereinzelt auf (4 Fälle erzielen 4 Maximalwerte der Stufe 4/5, 2 Fälle erzielen 5 Maximalwerte der Stufe 4/5, kein Fall erzielt 6 Maximalwerte der Stufe 4/5).

Für die USA-Studie ist zusammenfassend zu sagen, daß der Gesamtanteil von Postkonventionalität auf der Ebene der Issue- und der CJ-Werte zwar bei weitem nicht so hoch ist, wie es aufgrund der Verrechnungsregeln für die Stufenwerte möglich wäre. Der Anteil postkonventioneller Maximalwerte, die bei der Aggregation von CJ- zu Issue-Werten und von Issue- zu Form-Werten unterschlagen werden, ist aber durchaus bedeutsam.

Es wäre zu prüfen, ob in den anderen Studien, die auf der Grundlage des Standard Issue Scoring durchgeführt wurden, ähnliche Ergebnisse festgestellt werden können. Dabei wären auch die Verteilungen für die CJ-Werte differenziert zu analysieren und zu den Verteilungen für die Issue- und Form-Werte in Beziehung zu setzen.

Die für die USA-Studie ermittelten Ergebnisse lassen es zum einen fraglich erscheinen, ob die spezifische Form der Zusammenfassung der CJ-Werte zu Issue-Werten und der Issue-Werte zu Form-Werten mit Kohlbergs kompetenztheoretischen Ansprüchen vereinbar ist. Die Orientierung an der Konsolidierung von Kompetenzstrukturen, wie sie in den Aggregierungs-



regeln zum Ausdruck kommt, hat den Nachteil, daß den Maximalwerten nicht angemessen Rechnung getragen wird. Während das Interviewverfahren auf die Obergrenze der moralkognitiven Leistungsfähigkeit zielt, tendiert das Verfahren zur Ermittlung von aggregierten Stufenwerten dahin, diese oberste Leistungsgrenze zugunsten von Modalwerten zu unterschlagen. Unsere Analysen zeigen, daß beim Aggregationsprozeß Maximalwerte von Personen in nicht unerheblichem Ausmaß neutralisiert werden. Um den Maximalwerten besser gerecht zu werden, müßte sich die Bestimmung von globalen Kompetenzmaßen an schwächeren Kriterien orientieren. Wie die entsprechenden Kriterien festzulegen sind, ist eine schwierige Frage. Wir haben bis jetzt keine zufriedenstellende Antwort gefunden.

Die von uns ermittelten Ergebnisse lassen zum anderen Kohlbergs Bestimmung von Postkonventionalität fraglich erscheinen. Daß die Anteile postkonventioneller Stufen nicht nur bei den Form-Werten, sondern auch bei den Issue- und CJ-Werten insgesamt recht niedrig sind, ist auch Ausdruck einer problematischen theoretischen und operationalen Definition von Postkonventionalität bei Kohlberg. Wir haben in Kapitel 2.1 zu zeigen versucht, daß einige der von Kohlberg als konventionell eingestuften CJs im Manual postkonventionelles Denken zum Ausdruck bringen. Der relativ geringe Gesamtanteil postkonventioneller Stufen auf der Ebene der Issue- und Form-Werte ist also nicht nur auf die Neutralisierung von postkonventionellen CJ-Werten durch den Prozeß der Aggregation zurückzuführen, sondern auch auf eine problematische Konzeptualisierung postkonventionellen Denkens. Während die Stufenbestimmungen postkonventionellen Urteilens die Anzahl der kodierten postkonventionellen CJs schon stark einschränken, haben die Verrechnungsregeln zur Folge, daß der Anteil postkonventioneller Werte weiter reduziert wird. Um den postkonventionellen Urteilskompetenzen besser gerecht zu werden, ist es notwendig, Postkonventionalität neu zu bestimmen und den Anteil postkonventioneller CJs im Manual zu steigern. Wir teilen damit die Intentionen von Autoren wie Snarey und Shweder, gehen jedoch bei der Neubestimmung der postkonventionellen Stufen einen anderen Weg.

### 3.2.2 *Intraindividuelle Konsistenz der Stufenwerte*

Von Kohlberg und den an ihm orientierten Autoren wurde die Frage nach der intraindividuellen Konsistenz der Stufenwerte bislang kaum gestellt (vgl. Boyes & Walker 1988; Eckensberger, 1995). Allein für die CJ-Werte wurden Annahmen über die Stufenkonsistenz formuliert und empirische Analysen durchgeführt. Hypothesen und Untersuchungen zu den Issue-Werten fehlen (s. aber Teo, Becker & Edelstein, 1995).

Hinsichtlich der CJ-Werte geht Kohlberg von einem "dominant stage model" aus. Er nimmt an, daß bei der Urteilsbildung der Personen eine bestimmte Stufe vorherrscht und maximal zwei (benachbarte) Stufen auftreten. Tritt eine dritte Stufe auf und übersteigt deren Anteil nicht 10%, so wird das Auftreten dieser dritten Stufe als Meßfehler interpretiert (vgl. Colby & Kohlberg, 1983). Übersteigt der Anteil 10%, wird er als reale Inkonsistenz gewertet.

Von Kohlberg wird das "dominant stage model" indessen nicht präzise ausformuliert. So bleibt unklar, ob er von systematischen Veränderungen der Stufenhomogenität im Entwicklungsverlauf ausgeht.

Walker & Taylor (1991) haben das "dominant stage model" von Kohlberg präzisiert und unter entwicklungsdynamischen Gesichtspunkten weiterentwickelt, indem sie ein zyklisches Modell des Stufenübergangs mit einander abwechselnden Phasen von stärkerer Stufenhomogenität und Phasen von stärkerer Stufenvariabilität konzipierten und die Prozesse der Veränderung der Stufenkonsistenz längsschnittlich analysierten (vgl. Berkowitz und Keller, 1994; Keller, 1994).

Das von Kohlberg und Walker vertretene "dominant stage model" scheint durch die Daten gut bestätigt.

Für Kohlbergs Studie in den USA wird berichtet, daß ein relativ hoher Anteil der Urteilsbildung auf der Modalstufe liegt (durchschnittlich 68% für Form A, 72% für Form B, 69% für Form C). Auf den zwei am häufigsten gebrauchten (und) benachbarten Stufen liegen durchschnittlich sogar fast 100% (98% für Form A, 97% für Form B, 99% für Form C). Nur 9% der Interviews haben einen Anteil einer dritten Stufe, der über der Meßfehlergrenze von 10% liegt (vgl. Colby & Kohlberg, 1983; S. 36f.). In den anderen Studien, in denen die Stufenhomogenität analysiert worden ist (nämlich in der Israel-Studie von Kohlberg sowie in Studien, die in Taiwan und Neuguinea durchgeführt wurden), wurden ähnliche Ergebnisse erzielt (vgl. Boyes & Walker, 1988).

Auch Walker und Taylor finden in ihrer Studie zu Prozessen des Stufenübergangs, daß ein großer Teil der Urteilsbildung der Probanden auf der Modalstufe liegt. Zudem sehen sie durch ihre Befunde die Annahme bestätigt, daß sich die Stufenvariabilität im Sinne seines Modells von abwechselnden Phasen der Konsolidierung und Phasen stärkerer Stufeninkonsistenz verändert (vgl. Walker & Taylor, 1991).

Die Untersuchungen der Stufenkonsistenz von Personen zu einem Meßzeitpunkt, die Kohlberg und die an ihm orientierten Autoren durchgeführt haben, sowie die entwicklungsdynamische Konsistenzanalyse von Walker und Taylor teilen vor allem drei Beschränkungen:

- Bei der Bestimmung der Anteile der Urteilsbildung auf der Modalstufe und den benachbarten Stufen wird nicht zwischen reinen Stufen und Zwischenstufen differenziert, obwohl im Manual die CJs nach reinen Stufen und Zwischenstufen qualitativ unterschieden werden. Qualitativ eigenständige Zwischenstufen werden als Prozentanteile von reinen Stufen verrechnet.
- Die Analyse der intraindividuellen Konsistenz, die in Form einer Bestimmung der Häufigkeitsanteile von reinen Stufen durchgeführt wird, erstreckt sich nur auf die Argumentation der Person zu einer Form und nicht auf die Argumentation zu mehreren Formen.
- Auch die Aussagen über die Inkonsistenzen innerhalb einer Form bleiben beschränkt. So informieren Kohlberg und die an ihm orientierten Autoren nur über die Anteile von Interviews, bei denen mehr als zwei Stufen auftreten und nicht über die Anteile von Interviews mit nur einer oder nur zwei Stufen. Walker erfaßt neben dem Stufenanteil auf der dominanten Stufe den Stufenanteil unter und über der dominanten Stufe, ohne dabei aber das Ausmaß der Abweichung der Maximal- und Minimalwerte vom Modus zu berücksichtigen.

Dadurch, daß in den entsprechenden Konsistenzanalysen die Zwischenstufen zu Anteilen von reinen Stufen verrechnet werden, sind auch die für die intraindividuelle Konsistenz erzielten empirischen Ergebnisse problematisch.

Erforderlich ist also eine Analyse der CJ-Werte, durch die der qualitativen Eigenständigkeit der Zwischenstufen Rechnung getragen wird. Zugleich müssen Stufenstreuungen auch über mehrere Formen erfaßt und Muster von Inkonsistenzen differenziert erforscht werden.

Eine solche Untersuchung mußten wir aufgrund der mangelhaften Datenlage hinsichtlich der CJ-Werte zurückstellen. Wir sind stattdessen der Frage der Konsistenz der Issue-Werte nachgegangen, die wie schon erwähnt bisher gänzlich vernachlässigt wurde. Im Gegensatz zu Kohlbergs und Walkers Konsistenzuntersuchungen der CJ-Werte haben wir dabei auch Zwischenstufen berücksichtigt und die Stufenhomogenität über alle drei Formen hinweg untersucht. Uns interessierte vor allem, ob bei diesen aggregierten Werten noch stärkere

Stufenstreuungen auftreten und wie gut sich das "dominant stage model", das von Kohlberg und Walker zur Beschreibung der intraindividuellen Konsistenz der CJ-Werte formuliert wurde, für die Beschreibung der Streuung dieser Werte eignet.

Eine Analyse des Ausmaßes der Abweichungen der Minimal- und Maximalwerte der Probanden (vgl. Teo, Becker & Edelstein, 1995; Tab. 2) ergab, daß nur ein geringer Prozentanteil von Fällen eine starke Stufenkonsistenz aufweist: es finden sich keine Personen, die nur eine einzige reine Stufe oder Zwischenstufe verwenden, und der Anteil der Fälle mit einer Abweichung der Werte um eine halbe Stufe übersteigt kaum 10%. Ein größerer Anteil der Fälle dagegen besitzt eine starke Stufeninkonsistenz: So finden sich Fälle, die 5 oder mehr halbe Stufen verwenden (also eine Inkonsistenz von 2 Stufen oder mehr haben) bei den 10jährigen mit einem Anteil von 23%, bei den 13-14jährigen mit einem Anteil von 29% und bei den 16 - 18jährigen mit einem Anteil von 17%. Im Erwachsenenalter nimmt der Anteil der inkonsistenten Fälle allerdings ab und liegt bei etwa 10%.

Das relativ starke Ausmaß der Stufenvariabilität zeigt sich auch daran, daß Personen, die einen oder mehrere Issue-Werte der Stufe 1 haben, nicht nur Werte der Stufe 1/2 und 2, sondern häufig auch Werte der Stufe 2/3, manchmal sogar Werte der Stufe 3 oder 3/4 haben. Zugleich erzielen Personen, die einen oder mehrere Issue-Werte der Stufe 5 haben, nicht nur Werte der Stufe 4/5, sondern häufig auch Werte der Stufe 4, manchmal sogar Werte der Stufe 3/4, und Personen mit Werten der Stufe 4/5 erhalten nicht nur Werte der Stufe 4 und 3/4, sondern zum Teil auch Werte der Stufe 3.

So besetzt z.B. der Fall 24 beim 1. Meßzeitpunkt zweimal Stufe 1, sechsmal Stufe 1/2, einmal Stufe 2, zweimal Stufe 2/3, zweimal Stufe 3 und einmal Stufe 3/4. Der Fall 27 erhält beim 1. Meßzeitpunkt dreimal Stufe 1, keinmal Stufe 1/2, viermal Stufe 2, fünfmal Stufe 2/3, dreimal Stufe 3 und einmal Stufe 3/4. Und der Fall 5 weist beim 3. Meßzeitpunkt, dreimal Stufe 1, einmal Stufe 1/2, sechsmal Stufe 2, zweimal Stufe 2/3, einmal Stufe 3 und einmal Stufe 3/4 auf.

Was die Fälle mit einem oder mehreren Maximalwerten der Stufe 5 angeht, so sollen ebenfalls 3 Fälle zur Veranschaulichung einer recht starken Stufenvariabilität dienen: Fall 2 beim 5. Meßzeitpunkt, der viermal Stufe 5, neunmal Stufe 4/5, viermal Stufe 4 und einmal Stufe 3/4 erzielt; Fall 23 beim 5. Meßzeitpunkt, der dreimal Stufe 5, siebenmal Stufe 4/5 und sechsmal Stufe 4 erreicht; Fall 67 beim 6. Meßzeitpunkt, der zweimal Stufe 5, zweimal Stufe 4/5, zweimal Stufe 4 und zweimal Stufe 3/4 zugewiesen bekommt.

Zur Veranschaulichung einer recht starken Stufenvariabilität bei Fällen mit einem oder mehreren Maximalwerten der Stufe 4/5 seien folgende Fälle aufgeführt: Fall 65 beim 2. Meßzeitpunkt, der zweimal Stufe 4/5, einmal Stufe 4, einmal Stufe 3/4, zweimal Stufe 3, einmal Stufe 2/3 und zweimal Stufe 2 erzielt; Fall 17 beim 4. Meßzeitpunkt, der zweimal Stufe 4/5, zweimal Stufe 4, fünfmal Stufe 3/4, viermal Stufe 3 erzielt; Fall 25 beim 5. Meßzeitpunkt, der dreimal Stufe 4/5, siebenmal Stufe 4, zweimal Stufe 3/4 und fünfmal Stufe 3 erzielt.

Dabei ist zu bedenken, daß Probanden nur selten bei allen 18 Issues einen Stufenwert erhalten.

Die Konsistenzanalyse der Issue-Werte für alle Formen zeigen, daß eine Streuung der Werte über mehr als 2 Stufen häufiger auftritt. In Form einer reinen Stufe oder einer Zwischenstufe ist des öfteren eine dritte Stufe festzustellen. Bei Fällen mit einer Stufenvariation von 5 halben Stufen und mehr liegt notwendigerweise immer eine dritte Stufe vor. Der Anteil dieser Fälle ist nicht unerheblich. Aber auch bei allen Fällen, die eine Streuung über 4 halbe Stufen besitzen, wird noch eine dritte Stufe verwendet (wie z.B. dort, wo der Minimalwert Stufe 1/2 und der Maximalwert Stufe 3 beträgt).

Andererseits ist festzuhalten, daß ein großer Teil der individuellen Urteilsbildung auf dem Modus konzentriert ist (vgl. Teo, Becker & Edelstein, 1995; Tab. 3). In allen Altersgruppen liegen durchschnittlich etwa 50% der Urteilsbildung auf dem Modus. Relativ groß ist auch der Anteil der Urteilsbildung, der unmittelbar um dem Modus liegt. Durchschnittlich zwischen 30% und 40% der Urteile weichen eine halbe Stufe nach unten und nach oben vom Modus ab.

Eine Prüfung der Stufenverteilungen für die einzelnen Probanden ergibt jedoch, daß bei vielen Fällen keine bestimmte Stufe vorherrscht. So weisen auch einige relativ inkonsistente Fälle (Fälle mit 5 halben Stufen und mehr) eine stärkere Besetzung der Stufenrandbereiche auf.

Das "dominant stage model" von Kohlberg und Walker eignet sich also nur für einen Teil der Fälle zur Beschreibung der Konsistenz der Issue-Werte über alle Formen. Es treten häufig noch Streuungen über mehr als 2 Stufen auf und in nicht unerheblichem Ausmaß ist auch kein ausgeprägter Modus vorhanden.

### 3.2.3 Kontextabhängigkeit der Stufenwerte

Im Rahmen des Projekts haben wir uns die Frage gestellt, wieweit die intraindividuellen Stufenvariationen auf den Einfluß der einzelnen Issues, Fragetypen und Dilemmata zurückgeführt werden können.

Als mögliche Erklärungsfaktoren von Stufeninkonsistenzen innerhalb der Dilemmata kommen die verschiedenen konfligierenden Normen (Issues) in Betracht, zu denen der Proband Stellung beziehen muß, sowie die beiden unterschiedlichen Typen von Fragen, die zu diesen Normen gestellt werden, nämlich konzeptuelle und situationsspezifische Fragen. Bei der Analyse von Stufeninkonsistenzen über alle 3 Interviewformen sind neben den Issues und Fragetypen auch die unterschiedlichen Dilemmata als mögliche Einflußfaktoren relevant.

Kohlberg hat in der USA-Studie die mögliche Kontextabhängigkeit der Stufenwerte nicht untersucht. Er räumt allerdings ein, daß sich die Form-Werte für die 3 Formen unterscheiden. Die auftretenden Stufenunterschiede nimmt er dabei zum Anlaß, einen Korrekturfaktor einzuführen, mit dem die Parallelität der 3 Formen sichergestellt werden soll: "... There are small absolute differences between the forms. The differences affect the results of intervention studies with a pre- and posttesting design where one form is used as the pretest and the other as a posttest ... It can be seen that each form gives slightly different pulls for certain stages, for example, for stage 1 and for stage 5. Because of the slight difference in "pull" between the forms, it is appropriate to provide a conversion formula or scale score which would equalize the three forms. Such a conversion formula or scale score adds a constant to the score on each form and a weighting factor for each form based on its correlations with the scores for the forms combined" (Colby et al. 1983, S. 23). Offen bleibt dabei, wie die Unterschiede in den Form-Werten zu erklären bzw. welche Issues oder Dilemmata für diese Unterschiede verantwortlich sind.

Während es mittlerweile eine größere Anzahl von Studien gibt, die die Urteilsbildung zu den hypothetischen Dilemmata von Kohlbergs moralischem Interview mit der Urteilsbildung zu alltagsnäheren Dilemmata vergleichen (vgl. Krebs et al., 1991), liegen bislang nur wenige Studien vor, die innerhalb des Interviews von Kohlberg dem möglichen Einfluß von Kontextfaktoren nachgehen. Die Untersuchung der Wirkung von einzelnen Issues, Fragetypen und Dilemmata ist bislang weitgehend theoretisches Postulat geblieben (vgl. Rest, 1983).

Der Einfluß von Fragetypen und der Einfluß der Dilemmata wurde bisher überhaupt noch nicht untersucht. Lediglich in zwei Studien wurde auf der Ebene von Issue-Werten die

Wirkung von Issues analysiert, nämlich in einer Studie von Nisan und Kohlberg (1982) und in einer Studie von de Vries und Walker (1986).

In ihrer Längsschnittuntersuchung in der Türkei erforschten Nisan und Kohlberg unter anderem auch die Abhängigkeit der Stufenwerte von den 6 verschiedenen Issues einer Form (Leben, Gesetz, Gewissen, Strafe, Versprechen, Autorität). Sie nehmen an, daß bei den Issues "Gesetz" und "Strafe" niedrigere Werte erzielt werden als bei den anderen Issues: "The comparison between levels of judgment on different issues is stimulated by findings suggesting that the level of cognitive functioning in a certain domain (with regard to a certain content) may be affected by the individuals' familiarity with and experience in this domain (Flavell, 1977; Piaget, 1972). This leads to the hypotheses that when dealing with issues which are closer to their everyday experience, subjects might argue on a higher level than when reasoning about dilemmas which are far from their experience. This may introduce cultural variance in the level of moral judgment across different issues, due to differences in life conditions and environmental demands (Berry, 1971). Ethnographic material suggested that the issues of law and punishment (by law) are further from our subjects' experience than those of (a father's) authority, contract (keeping promises) and life" (Nisan & Kohlberg, 1982, S. 867). Nisan und Kohlberg überprüften ihre Annahme an einer Population von 10 bis 28 Jahre alten Probanden. Sie fanden, daß die Werte bei den Issues "Gesetz" und "Strafe" in der Tat niedriger liegen als bei den anderen Issues, aber dies allein bei der Altersgruppe der 10 bis 15jährigen. Bei der zusätzlich erfaßten Altersgruppe der 16 bis 28jährigen waren signifikante Unterschiede nicht festzustellen. Nisan und Kohlberg präsentieren dabei lediglich Ergebnisse für die 6 Issues der Form A.

Anhand einer Population von 17 bis 35 Jahre alten amerikanischen Universitätsstudenten untersuchen de Vries und Walker (1988) unter anderem auch den Einfluß der Issues auf die Stufenwerte. Ihre Studie bezieht sich allerdings nur auf die Urteilsbildung zu den Issues "Gewissen" und "Strafe" und beschränkt sich darüber hinaus auf Stellungnahmen zu konzeptuellen Fragen: Die Probanden hatten sich in Form eines schriftlichen Aufsatzes mit dem Für und Wider der Todesstrafe auseinanderzusetzen und damit Fragen zu beantworten, die Bestandteil des Fragenkatalogs zum "Dr. Rogers-Dilemma" sind. De Vries und Walker stellen fest, daß nicht die Entscheidungsrichtung, sondern das Issue den stärksten Einfluß auf die Stufenwerte hat: viele Befürworter der Todesstrafe haben wie die Gegner der Todesstrafe bei der Argumentation gegen diese Strafe (Issue "Gewissen") höhere Werte als bei der Argumentation für die Todesstrafe (Issue "Strafe"). Die Ergebnisse werden von de Vries und Walker aber nicht in differenzierter Form berichtet, und es wird auch keine Erklärung für die ermittelte Issueabhängigkeit der Urteilsbildung geliefert.

Nisan und Kohlberg sowie de Vries und Walker liefern also keine systematische Analyse des Issue-Einflusses. Ihre Untersuchung erstreckt sich nicht auf alle 9 Dilemmata, womit sich die Frage nach der Generalisierbarkeit der jeweils ermittelten Befunde aufdrängt. Während die Studie von de Vries und Walker die Erwartung systematischer Issue-Effekte im Erwachsenenalter nahelegt, vermittelt die Studie von Nisan und Kohlberg den Eindruck, daß Issue-Effekte nur in der frühen und mittleren Adoleszenz zu finden sind.

Hinsichtlich der Frage der Dilemmaabhängigkeit der Stufenwerte ist eine Untersuchung von Eckensberger (1989) bedeutsam. Darin werden zwar nicht Kohlbergs moralisches Interview und das Standard Issue Scoring verwendet, sondern modifizierte Verfahren. Doch benutzt Eckensberger die 3 Dilemmata der Form A, nämlich "Heinz", "Officer Brown" und "Joe" (das Joe-Dilemma wurde dabei leicht verändert). Seiner Untersuchung des Dilemmaeinflusses liegt die Differenzierung zwischen "personalen", "transpersonalen" und "gemischten" Dilemmata zugrunde: "Some dilemmas (personal dilemmas) deal with conflicts that exist primarily between acting subjects in terms of persons, others (transpersonal dilemmas) primarily deal with functions, roles or general principles, and still others are more or less indifferent with reference to these social spaces (equivocal dilemmas)" (Eckensberger, 1988, S. 2). Eckensberger klassifiziert das Joe-Dilemma als personales Dilemma, "Officer Brown" als transpersonales Dilemma und das Heinz-Dilemma als gemischtes Dilemma. Bei der von ihm untersuchten Population von 10 bis 30 Jahre alten Personen findet er, daß ab dem Alter von 16 Jahren bei "Joe" im Durchschnitt niedrigere Stufenwerte erzielt werden als bei "Officer Brown", während die Werte bei "Heinz" höher liegen als bei "Joe" und niedriger als bei "Officer Brown".

Es drängt sich die Frage auf, ob auch in der USA-Studie ähnliche Dilemmaeffekte bei Form A auftreten und ob Kontexteinflüsse in den anderen Formen zu finden sind.

Das Projekt bemühte sich um eine systematische Analyse des Einflusses der Issues, Dilemmata und Fragetypen. Die Untersuchung bezieht sich auf alle Issues und Dilemmata des moralischen Interviews.

Zur Abhängigkeit der Urteilsbildung von Issues und Dilemmata liegen erste Ergebnisse vor und zwar hinsichtlich der Issue-Werte. (Eine Analyse des Einflusses der Fragetypen konnte noch nicht vorgenommen werden. Dies setzt die CJ-Werte als empirische Grundlage voraus).



Der Untersuchung der Issueabhängigkeit von Stufenwerten liegt die Differenzierung zwischen heteronomen und autonomen Issues zugrunde. Die Gruppe der heteronomen Issues ("Gesetz", "Strafe", "Autorität") umfaßt Begründungen der Bedeutung der Konformität mit bestehenden Institutionen und mit Autoritätspersonen. Die Gruppe der autonomen Issues ("Leben", "Gewissen", "Versprechen") beinhaltet Begründungen der Bedeutung der individuellen Rechte und von Reziprozitätsnormen. Wir nehmen an, daß auf präkonventionellem Niveau bei heteronomen Issues niedrigere Werte erzielt werden als bei autonomen Issues, daß hingegen auf konventionellem Niveau bei heteronomen Issues höhere Werte erreicht werden als bei autonomen Issues: Für präkonventionelle Probanden erwarten wir, daß diese bei der Begründung heteronomer Issues (z.B. bei Antworten auf die Frage, warum es wichtig ist, dem Gesetz zu folgen) vorwiegend auf die Forderungen und Sanktionen von Autoritäten Bezug nehmen, was bei Kohlberg als Stufe 1 kodiert wird. Bei der Begründung autonomer Issues (z.B. bei Antworten auf die Frage, warum es wichtig ist, menschliches Leben zu retten), sollten die präkonventionellen Probanden eher die Ansprüche und Eigeninteressen von Personen geltend machen, was als Stufe 2 kodiert wird. Hinsichtlich der konventionellen Probanden vermuten wir, daß die für Stufe 4 kennzeichnende Orientierung an den Funktionsbedingungen sozialer Systeme bei der Begründung heteronomer Issues deutlich stärker auftritt als bei der Begründung autonomer Issues. Im Gegensatz zu Nisan und Kohlberg gehen wir also einerseits von einer Niveauabhängigkeit der Issue-Effekte aus, andererseits führen wir die erwarteten Effekte nicht auf Unterschiede in der Vertrautheit mit den Issues zurück, sondern auf eine durch die Issues nahegelegte Präferenz stufentypischer Begründungsgesichtspunkte.

Bei der Untersuchung des Dilemmaeinflusses unterscheiden wir in Anlehnung an Eckensberger zwischen personalen, transpersonalen und gemischten Dilemmata. Als personale Dilemmata lassen sich in Form A "Joe", in Form B "Judy" und in Form C "Bob, Karl" klassifizieren. Als transpersonale Dilemmata lassen sich in Form A "Officer Brown", in Form B "Dr. Rogers" und in Form C "Korea" einordnen. Als gemischte Dilemmata können in Form A "Heinz", in Form B das Euthanasie-Dilemma und in Form C "Valjean" betrachtet werden. Ähnlich wie Eckensberger vermuten wir, daß auf konventionellem Niveau bei personalen Dilemmata vorwiegend eine für Stufe 3 kennzeichnende Beziehungsperspektive verwendet wird, bei transpersonalen Dilemmata hingegen vorwiegend eine für Stufe 4 kennzeichnende Systemperspektive. Bei gemischten Dilemmata sollten die Stufenwerte durchschnittlich höher liegen als bei personalen Dilemmata und niedriger als bei transpersonalen Dilemmata.

Wir haben diese Annahmen hinsichtlich des Issue- und Dilemmaeinflusses durch eine globale Analyse statistisch überprüft.

Um den Einfluß von heteronomen und autonomen Issues zu untersuchen, haben wir bei jeder Altersgruppe den Mittelwert von allen heteronomen und den Mittelwert von allen autonomen Issue-Werten errechnet. Dabei zeigt sich, daß bei der Altersgruppe der 10jährigen sowie bei der Altersgruppe der 13- und 14jährigen - in der jeweils präkonventionelle Stufen vorherrschen - der Mittelwert für die autonomen Issues wie erwartet signifikant höher liegt als der Mittelwert für die heteronomen Issues. Hingegen treten in den späteren Altersgruppen im Widerspruch zu unserer Erwartungen keine signifikanten Unterschiede bei den Mittelwerten mehr auf (vgl. Teo, Becker & Edelstein, 1995).

Der Einfluß des personalen und des transpersonalen Dilemmatyps wurde in der Weise untersucht, daß (entsprechend der Analyse von Issue-Effekten) bei jeder Altersgruppe der Mittelwert von allen personalen Issue-Werten und der Mittelwert von allen transpersonalen Issue-Werten ermittelt wurde. Wir fanden, daß bei der Altersgruppe der 16- bis 18jährigen, in der die Stufen 3/4 und 4 erstmals stärker auftreten, der Mittelwert für die transpersonalen Dilemmata signifikant höher liegt als der Mittelwert für die personalen Dilemmata. Auch bei den Altersgruppen der 20- bis 22jährigen, der 24- bis 26jährigen und der 28- bis 30jährigen zeigen sich signifikante Unterschiede im Sinne unserer Erwartungen. Bei den letzten beiden Altersgruppen treten allerdings keine signifikanten Unterschiede mehr auf (vgl. Teo, Becker & Edelstein, 1995, Tab. 4).

Die Untersuchung der Kontextabhängigkeit der Stufenwerte auf der Ebene einzelner Issues und Dilemmata ist noch nicht abgeschlossen. Eine erste Analyse der entsprechenden Issue-Werte zeigte, daß die auf der globalen Ebene ermittelten Effekte mit Einschränkungen auch auf der Ebene von einzelnen Issues und Dilemmata zu finden sind.

So liegen im Alter von 16 bis 30 Jahren die Mittelwerte für die beiden transpersonalen Issues zu "Officer Brown", zu "Dr. Rogers" und zu "Korea" im allgemeinen höher als die Mittelwerte für die beiden personalen Issues zu "Judy" und zu "Bob, Karl". Allerdings liegen nur die Mittelwerte für die Issues zu "Officer Brown" höher als die Mittelwerte für die Issues zu "Joe"; die Werte zu "Dr. Rogers" und "Korea" unterscheiden sich kaum von den Werten im Joe-Dilemma. Hinsichtlich dieses Dilemmas scheinen sich unsere Erwartungen also nur teilweise zu bestätigen.

Darüber hinaus ist festzustellen, daß die gemischten Dilemmata, nämlich "Heinz", "Euthanasie" und "Valjean", häufig ähnlich hohe Mittelwerte bei den Issues erzielen wie die transpersonalen Dilemmata. Die Differenzierung zwischen transpersonalen und gemischten Dilemmata scheint also empirisch nicht bedeutungsvoll zu sein.

Hinsichtlich der Werte für die heteronomen und autonomen Issues eines Dilemmas zeigt sich, daß bei der Altersgruppe der 10jährigen mit der Ausnahme nur eines Dilemmas (nämlich "Korea") der Mittelwert für das heteronome Issue eines Dilemmas niedriger liegt als der Mittelwert für das autonome Issue. Bei der Altersgruppe der 13-14jährigen liegen allerdings bei 3 Dilemmata, nämlich bei "Joe", "Euthanasie" und "Dr. Rogers" die Mittelwerte für das heteronome Issue nicht niedriger als für das autonome Issue.

Bei den Altersgruppen, bei denen unsere globale Analyse von Kontexteinflüssen keine signifikanten Unterschiede erbracht hat, ist auch keine Systematik in der Variation von Issue-Werten zu einzelnen Issues und Dilemmata zu entdecken. Zu erwähnen ist lediglich, daß zwischen 16 und 30 Jahren beim Issue "Strafe" zu "Officer Brown" gewöhnlich der höchste Mittelwert erzielt wird.

Zusammenfassend ist hinsichtlich des Einflusses der Dilemmata zu sagen, daß die von Eckensberger eingeführte und von uns übernommene Differenzierung zwischen personalen, transpersonalen und gemischten Dilemmata empirisch nur zum Teil als sinnvoll bestätigt. Sowohl in Form A als auch in Form B und C erreichten transpersonale und gemischte Dilemmata häufig ähnliche Werte. Es scheint lediglich eine Differenzierung zwischen personalen und nicht-personalen Dilemmata bedeutsam. Diese Unterscheidung erscheint freilich theoretisch ausgesprochen sinnvoll und verweist auf eine ethische Distinktion, die weitere Analysen rechtfertigt.

Hinsichtlich des Einflusses von Issues ist zusammenfassend zu sagen, daß es ähnlich wie in der Studie von Nisan und Kohlberg nur in der frühen und mittleren Adoleszenz zu systematischen Issue-Effekten kommt. Die durch Walkers Ergebnisse nahegelegte und von Nisan und von uns gehegte Erwartung, daß systematische Issue-Effekte im Erwachsenenalter auftreten, ließ sich nicht bestätigen.

In der USA-Studie erklärten also die Issues und Dilemmata nur zum Teil die doch recht beträchtlichen intraindividuellen Stufenvariationen. So bleibt die Frage unbeantwortet, wie der Teil der Stufeninkonsistenzen, der nicht auf den Einfluß von Issues und Dilemmata zurückzuführen ist, erklärt werden kann.

### 3.3 Überprüfung der Interraterreliabilität für die Criterion Judgments

Wir haben im Abschnitt 3.1 kritisiert, daß bei der Analyse der Daten, die mit Hilfe des Standard Issue Scoring und des Verfahrens zur Kodierung von Typen ermittelt wurden, die niedrig aggregierten Werte kaum Berücksichtigung fanden. In Abschnitt 3.2 haben wir dann einige Ergebnisse unserer Untersuchung der niedrig aggregierten Issue-Werte vorgestellt.

Für die von uns eingeklagte und zum Teil durchgeführte Mikroanalyse moralischen Urteilens ist eine hohe Reliabilität bei den niedrig aggregierten Werten notwendig. Insbesondere eine hohe Interraterreliabilität, d.h. große Übereinstimmung in den Ergebnissen von Kodierungen der gleichen Interviews durch verschiedene Rater ist erforderlich.

Voraussetzung für eine Mikroanalyse der Stufenentwicklung ist eine hohe Interraterreliabilität bei der Kodierung der Issue- und CJ-Werte. Die Untersuchung von Normen und Elementen im Rahmen des Standard Issue Scoring sowie von Typenkriterien und dilemmabezogenen Typen-Werten verlangt ebenfalls eine reliable Kodierung auf niedrigem Aggregationsniveau.

Die Frage, ob im Rahmen des Standard Issue Scoring und des Typenscoring auf niedrigem Aggregationsniveau reliabel kodiert werden kann, stellt sich aber auch unabhängig von mikroanalytischen Absichten. Um die Güte beider Auswertungsverfahren einzuschätzen, reichen Reliabilitätsaussagen hinsichtlich der Form-Werte nicht aus. Hohe Reliabilitäten bei den hoch aggregierten Werten implizieren keineswegs hohe Reliabilitäten bei den niedrig aggregierten Werten.

Im Hinblick auf das Standard Issue Scoring ergibt sich die Notwendigkeit einer Untersuchung der Interraterreliabilität für die niedrig aggregierten Werte auch aus Kohlbergs eigenen Ansprüchen. Kohlberg verknüpft mit dem Standard Issue Scoring den Anspruch, daß dieses Verfahren eine wesentlich reliablere Messung von moralischen Urteilen erlaubt als das frühere Structural Issue Scoring. Dadurch, daß sich die Kodierung an kleinen Auswertungseinheiten orientiert und präzise Kriterien für die Zuordnung von Stufenwerten bzw. Inhalten zu Intervieweinheiten vorgegeben werden, soll die subjektive Willkür im Prozeß des Kodierens eingeschränkt werden. Damit einer Aussage ein Criterion Judgment zugewiesen wird, müssen spezifische Kriterien erfüllt sein, die im Manual aufgeführt sind: "A structural or theoretical explication has been provided for each criterion judgment in the manual. Each explication includes a statement of the underlying stage structure reflected in

the criterion judgment (Stage Structure), detailed criteria for defining a match to the criterion judgment (Critical Indicators), explanations of distinctions among criterion judgments a scorer is likely to confuse, and several examples of interview material which can be considered to match the criterion judgment ...This criterion judgment format was an important advance over earlier scoring systems. The subjectivity of scoring decisions is minimized by the concrete and explicit specification of exactly what constitutes a match between interview material and a criterion judgment in the scoring manual" (Colby, Kohlberg 1983, S. 12).

Entsprechend sehen die operationalen Kriterien für das CJ#13 des Issues life beim Heinz-Dilemma wie folgt aus:

Criterion Judgment: 1. [Heinz should steal the drug, OR it is Heinz's duty to steal the drug] because he should care about her, or love her; or because he is supposed to stick close to her; OR because he shouldn't just sit back and watch her die.  
2. [Heinz should steal the drug even if he doesn't love his wife] because he must have loved her at one time or should still care about her; OR because he is still her husband.

Stage structure: As in Criterion Judgment #12, stealing is justified insofar as it conforms in this case to a norm of affiliation and concern between husband and wife or friends. There is a focus in this criterion judgment on obligation as based on shared expectations for prosocial role behavior.

Critical indicators: Required for a match is an appeal to either: (a) how Heinz should feel (should care about her, love her, must have loved her once, should still care about her, etc.); or (b) proper empathic conduct (shouldn't just sit back and let her die).

Distinctions and inclusions:

Within stage 3: Do not confuse with alternative Stage 3 points which focus on Heinz's prosocial character motives in general (CJ#16), or on the concern which Heinz *would* feel (CJ#12). Between other stages: Do not confuse this prescription for Heinz to care with the ambiguous 3/4 ideal of the responsibility Heinz would feel (CJ#21). Distinguish from the Stage 4 understanding of role norms in terms of societal responsibilities or contractual obligations (CJ#27). This point should be included within the cited Stage 3/4 or Stage 4 judgments.

Match examples: 1. Is there a good reason to steal if Heinz does not love his wife? - Yes, just to try to save a life. He, at some time, must have loved her and he is still her husband.

2. Is it a husband's duty to steal the drug for his wife if she can get it no other way? Would a good husband do it? I don't think any husband should sit back and watch his wife die. It is his duty to save her life.

Es stellt sich die Frage, wie gut es Kohlberg gelungen ist, durch strenge operationale Kriterien die subjektive Willkür im Prozeß des Kodierens einzuschränken.

Von Kohlberg wurde die Interraterreliabilität des Standard Issue Scoring und des Typenscoring ausschließlich auf der Ebene der Form-Werte überprüft. Dabei ergaben sich jeweils relativ hohe Reliabilitäten. Für die globalen Stufenwerte zu den einzelnen Formen A, B und C berichtet Kohlberg von einer Reliabilität zwischen 75% und 88% auf der 9-Stufenskala (mit 5 reinen Stufen und 4 Zwischenstufen). Für die Form-Werte im Rahmen des Typenscoring stellt er eine prozentuale Übereinstimmung zwischen 80% und 100% bei der Kodierung von heteronomer und autonomer Orientierung fest.

Auch von der an Kohlberg orientierten Forschung wurden Resultate zur Interraterreliabilität gewöhnlich nur für die Form-Werte ermittelt. Dabei ergaben sich für die Stufen (vgl. z.B. Walker, 1988) und für die Typen (vgl. z.B. Logan, Snarey, Schrader; 1990) ähnlich hohe Reliabilitäten wie bei Kohlberg.

Die Frage, inwieweit sich die Werte zu den einzelnen Typenkriterien und die Typenwerte für die einzelnen Dilemmata reliabel kodieren lassen, hat die an Kohlberg orientierte Forschung bislang überhaupt nicht gestellt. Im Hinblick auf das Standard Issue Scoring informieren lediglich Walker (1989) und Pratt et al. (1984) über Interraterreliabilitäten bei der Kodierung von Elementen. Walker berichtet von einer Reliabilität von .80 und Pratt et al. von Reliabilitätswerten für die 4 Elementegruppen, die zwischen .54 und .92 liegen. Beide Autoren prüfen nur, ob sich die 4 Elementegruppen reliabel kodieren lassen. Sie erfassen nicht die Reliabilität für die Kodierung der einzelnen Elemente. Ergebnisse für die Kodierung von Normen liegen ebensowenig vor wie Resultate für die Kodierung von Stufen auf der Ebene von Issue- und CJ-Werten.

Bislang ist also nicht geklärt, ob die Interraterreliabilitäten für die niedrig aggregierten Werte ähnlich hoch liegen wie für die hoch aggregierten Form-Werte. Die Frage, inwieweit eine Mikroanalyse der Stufen und Inhalte möglich ist, bleibt unbeantwortet. Über die Güte des Typenscoring läßt sich bisher ebensowenig eine zureichende Aussage machen wie über die Güte des Standard Issue Scoring. Offen ist auch, in welchem Maße die strengen operationalen

Bedingungen des Standard Issue Scoring zu einer Einschränkung der subjektiven Willkür im Prozeß des Kodierens führen.

Kohlberg und die an Kohlberg orientierte Forschung haben darüber hinaus bisher auch nicht hinreichend geprüft, wie stark die erzielten Reliabilitätswerte vom Ausmaß der Kodiererfahrung der Rater abhängen.

Mit dem Standard Issue Scoring verknüpft Kohlberg den Anspruch, daß auch relativ unerfahrene Rater das Verfahren reliabel anwenden können. Er meint nachgewiesen zu haben, daß unerfahrene Rater hohe Reliabilität erzielen: "Raters who scored these reliability data varied in degree of experience using the manual. Of the five form A raters two were familiar with Structural Standard Issue Scoring as well as Standard Issue Scoring and in fact were the authors of Standard Issue forms A and B (Colby and Gibbs). The third rater was highly experienced but was not the author of the manual. The other two less experienced raters had learned to score within the prior 6-8 months using the manual. They consulted with more experienced raters as they learned to use the manual to score a separate set of practice cases, but this consultation was no more extensive than would be provided at the week-long training workshop held at Harvard each year. Interrater reliability figures between the two "new scorers" and between each of them and an experienced scorer were at least as high as reliability figures among the experienced scorers. The same was true for form B interrater reliability, which involved two experienced and two new scorers. Thus it seems warranted to conclude that the Standard Issue Manual can be reliably mastered by relatively inexperienced users" (Colby & Kohlberg, 1983; S. 20f.)

Es ist aber problematisch, Rater mit einer Kodierpraxis von 6 - 8 Monaten als unerfahrene Rater zu bezeichnen. Zudem bedeutet eine hohe Übereinstimmung, die zwischen mehr oder weniger erfahrenen Ratern bei der Kodierung von Form-Werten erzielt wird, nicht unbedingt eine hohe Übereinstimmung bei der Kodierung von CJ- oder Issue-Werten. Kohlbergs Behauptung, daß auch relativ unerfahrene Rater mit dem Standard Issue Scoring Manual reliabel umgehen können, ist also nicht hinreichend belegt. Um zu klären, in welcher Weise die erzielten Reliabilitäten von der Kodiererfahrung abhängen, wären auch Rater mit wesentlich weniger Kodierpraxis heranzuziehen und Unterschiede in der Reliabilität der Kodierung von Werten auf verschiedenen Aggregationsniveaus zu berücksichtigen.

Im Hinblick auf das Typenscoring wurde weder von Kohlberg noch von der an Kohlberg orientierten Forschung die Frage nach dem Einfluß der Kodiererfahrung auf die Interraterreliabilität aufgeworfen.

Aber nicht nur Informationen über die Interraterreliabilität für die niedrig aggregierten Werte und Resultate zum Einfluß der Kodiererfahrung auf die Interraterreliabilität fehlen. Auch nach den möglichen Ursachen für die Abweichung von Kodierergebnissen ist bisher nicht gefragt worden. So ist ungeklärt, ob die einzelnen Kodierschritte in unterschiedlichem Maße für die Einschränkung der Reliabilität verantwortlich sind.

Das Standard Issue Scoring und das Verfahren zur Kodierung von Typen besteht jeweils aus 3 Kodierphasen. Bei beiden Verfahren hat der Rater zunächst eine Entscheidung darüber zu treffen, welche Aussagen des Probanden zu einem Dilemma überhaupt moralische Urteile darstellen. Die Aussagen müssen 3 Kriterien genügen, um als kodierbare moralische Urteile zu zählen. Es dürfen (a) nur Stellungnahmen berücksichtigt werden, die präskriptiven Charakter haben. Zugleich müssen (b) die präskriptiven Aussagen eine Rechtfertigung für die Verbindlichkeit von Handlungen oder Werten liefern und nicht bloß Präferenzen des Probanden zum Ausdruck bringen. Es sind (c) nur solche Begründungen heranzuziehen, die die Person als stichhaltig betrachtet, die "subjektiv valide" sind.

In der zweiten Kodierphase bei beiden Verfahren hat der Rater zu entscheiden, wieweit die im Manual vorgegebenen Kriterien von den als moralisch klassifizierten Aussagen erfüllt werden. Im Rahmen des Standard Issue Scoring ist zu prüfen, ob bestimmten Intervieweinheiten (IJs) eines Dilemmas bestimmte CJs im Manual entsprechen. "You must ... determine whether the correspondence between the IJ and the CJ is substantial enough to warrant its being scored as a match. That is, you must decide if the IJ is really an example of the kind of thinking demonstrated in the CJ" (Colby & Kohlberg, 1987; S. 172). Im Rahmen des Typenscoring ist die Argumentation des Probanden zu einem Dilemma daraufhin zu untersuchen, ob sie bei den dilemmaspezifisch operationalisierten Typenkriterien Definitionsmerkmale des autonomen oder Definitionsmerkmale des heteronomen Typus aufweist: " ... we have developed idealized forms of each of the nine criteria for each dilemma: a "B" form that represents the idealized autonomous judgment, and an "A" form that represents the idealized heteronomous judgment. The scorer's task is to determine which of the two forms of the criteria in question a given subject's moral judgment most closely resembles" (Colby & Kohlberg, 1987, S. 353f.).

In der dritten Kodierphase hat dann der Rater die auf der Ebene der CJs ermittelten Kodierungen bzw. die Werte für die einzelnen Typenkriterien nach bestimmten Verrechnungsregeln zu globalen Werten zusammenfassen. Beim Standard Issue Scoring sind Stufenwerte für die CJs zu Issue-Werten und Issue-Werte zu Form-Werten zu aggregieren, beim Typenscoring sind



die Werte von einzelnen Typenkriterien zu Dilemma-Werten und die Dilemma-Werte zu Form-Werten zusammenzufassen.

Eine Abweichung in den Ergebnissen der Kodierung von CJs bzw. Typenkriterien kann deshalb zum ersten dadurch zustande kommen, daß sich die Rater bei ihrer Einschätzung unterscheiden, ob eine bestimmte Intervieweinheit für die Kodierung relevant ist oder nicht. Während ein Rater eine bestimmte Aussage als moralisch gehaltvoll interpretiert und die Stellungnahme kodiert oder bei der Kodierung von anderen Aussagen berücksichtigt, läßt ein anderer Rater diese Aussage bei der Kodierung außer Acht, weil sie in seinen Augen eines oder mehrere der 3 Kriterien für ein moralisches Urteil nicht erfüllt.

Eine Abweichung in den Kodierergebnissen für die CJs bzw. die Typenkriterien kann zum zweiten aus unterschiedlichen Einschätzungen der Rater bei der zweiten Kodierphase resultieren: die Rater klassifizieren eine bestimmte Aussage zwar gemeinsam als moralische Stellungnahme, unterscheiden sich aber bei ihrer Bewertung, ob diese Aussage der im Manual angegebenen prototypischen Argumentation entspricht.

Aber auch der Prozeß der Aggregierung von Werten kann für Abweichungen in den Kodierergebnissen verantwortlich sein. Fehler bei der Aggregierung mindern die Reliabilität der höher aggregierten Werte.

Der dritte zentrale Kodierschritt in jedem Verfahren trägt wahrscheinlich am wenigsten zur Einschränkung der Interraterreliabilität bei. Die vorzunehmende Aggregierung von Werten verlangt jeweils nur eine mechanische Anwendung der Verrechnungsregeln. Obwohl besonders beim Standard Issue Scoring die Regeln für die Aggregierung von Werten relativ komplex sind, dürften Rater nach einiger Kodierpraxis keine Fehler bei der Aggregierung mehr machen.

Die erste und zweite Kodierphase in jedem Verfahren führt dagegen vermutlich weit stärker zur Einschränkung der Interraterreliabilität. Diese Phasen erfordern jeweils relativ komplexe Prozesse der Bewertung von Aussagen.

Beim zweiten zentralen Kodierschritt im Rahmen des Standard Issue Scoring soll, wie schon erwähnt, durch operationale Kriterien für die Zuordnung eines CJ zu einer Intervieweinheit eine Übereinstimmung in den Entscheidungen der Rater sichergestellt werden. Dadurch daß die Rater bei ihren Kodierentscheidungen einerseits zu prüfen haben, ob die das CJ definierenden "Critical Indicators" erfüllt sind, andererseits zu untersuchen haben, ob die

entsprechende Aussage mit der "Stage Structure" übereinstimmt, wird nicht nur eine hohe Reliabilität, sondern auch eine hohe Validität der Kodierung zu erreichen versucht: "... in order for a IJ-CJ comparison to be considered a match, (1) the IJ must exhibit the surface features of the CJ in question, and (2) the CJ must be consistent with the stage structure of that CJ. Our requirement that the IJ exhibit the surface features is designed to prevent an excessively loose interpretation in which the scorer attributes to the subject ideas that the subject may not in fact have been expressing. Our requirement that the IJ also be consistent with the unreeling moral judgment structure of the stage represented by the CJ is an attempt to minimize the invalidity of scores based on a literal matching of surface features of the IJ and the CJ. A scoring system based on a very literal "strict interpretation" approach might yield greater reliability, but responses could be misscored from the structural point of view. The increase in reliability would be outweighed by a decrease in validity. The adoption of a very structural orientation or "loose interpretation" approach might increase the validity of incidence of correct classification for very experienced scorers, but would involve greater unreliability in general as well as less valid scoring among less experienced raters" (Colby & Kohlberg, 1987; S. 172f.). Damit der Rater einer Intervieweinheit ein CJ zuordnen kann, müssen alle angegebenen critical indicators des CJ und muß zugleich auch die stage structure erfüllt sein.

Mit der Prüfung der stage structure ist aber die strukturelle Bewertung der Intervieweinheit noch nicht abgeschlossen. Der Rater hat zusätzlich zu prüfen, ob zwei weitere im Manual vorgegebene Kriterien erfüllt werden, nämlich die "Distinctions" und "Inclusions" eines CJ: "The IJ must not only be consistent with the CJ stage structure, it must also fail to provide a better match to another CJ. To make this decision, the Distinctions and Inclusions paragraph(s) of each explicated CJ should be examined. In the Distinctions paragraph, the CJ in question is distinguished from parallel or similar CJ ideas at the same and different moral stages. The Distinctions notes are designed to further clarify stage structure and to prevent misscoring an IJ as a match to a particular CJ when in fact it would be a better match to a similar-sounding - although conceptually distinct - CJ found elsewhere (either a CJ at a different stage or a CJ at the same stage but with a different norm and/or element) ... Included with the distinctions are notes on possible inclusions with other CJs ... Whenever interview material matches two CJs of the same norm and element but different stages, the lower stage CJ is considered to be included within the higher stage CJ. Only the score of the highest match on that norm and element is included when assigning a score for the issue. This same inclusion relationship sometimes holds for CJs with different elements but the same norm" (Colby & Kohlberg, 1987; S. 177f.).

In der zweiten Kodierphase des Standard Issue Scoring können also unterschiedliche Einschätzungen der Rater hinsichtlich der critical indicators, der stage structures, der distinctions und der inclusions für abweichende Ergebnisse bei der Kodierung von CJs verantwortlich sein.

Beim Standard Issue Scoring können Rater aber nicht nur in der Vergabe von CJs voneinander abweichen, sondern auch in ihrer Einschätzung, ob Guess- Werte zu vergeben sind bzw. welche Guess-Werte zuzuweisen sind. Guess-Werte hat der Rater dann zu ermitteln, wenn bei einem Issue eines Dilemmas kein einziges CJ vergeben werden kann: "Sometimes an interview will not yield even a single match ... for the issue being scored. In such cases, if any moral judgment material is available on the issue in question, the scorer must assign a guess score to the issue, however ambiguous the material on the issue might be ... We have introduced the concept of a guess score in order to (a) increase the likelihood of each subject's receiving a score for each of the six standard issues, thus maximizing standardization in sampling across the domain of moral judgment; (b) make possible the scoring of novel responses not represented in the manual; (c) minimize the introduction of experimental bias; the use of guess scores should minimize any tendency to eliminate as unscorable data that violate the scorer's expectations; and (d) compensate for the fact that raters often disagree as to the point at which material should no longer be considered stage scorable" (Colby & Kohlberg, 1987; S. 181f.). Guess-Werte können auf der Grundlage von CJs vergeben werden, wobei weniger strenge Zuordnungskriterien angewandt werden. Sie können aber auch durch eine intuitive Gesamteinschätzung des Interviewmaterials, die sich auf die allgemeine Stufendefinition stützt, ermittelt werden. Bei der Kodierung von Guess-Werten kann der Rater dabei auf Stellungnahmen zurückgreifen, die keinen präskriptiven Charakter besitzen.

Bei der Vergabe von Guess-Werten stellt insbesondere das zum Teil heterogene und mehrdeutige Interviewmaterial den Rater vor Schwierigkeiten. Unterschiedlichen Kodierresultate können z.B. dadurch zustande kommen, daß der eine Rater den Guess-Wert für ein Issue auf der Grundlage eines bestimmten CJs ermittelt hat, der andere Rater zu diesem Wert dagegen auf der Grundlage einer intuitiven Gesamteinschätzung der Äußerungen des Probanden kommt.

Die zweite Phase des Standard Issue Scoring verlangt also viele und z.T. recht schwierige Entscheidungen. Beschränkungen der Interraterreliabilität scheinen sich besonders hier zu ergeben. Kohlberg selbst stellt fest, daß dies die Phase ist, wo die stärksten Reliabilitätseinbußen auftreten (vgl. Colby & Kohlberg 1987; S. 172).

In der zweiten Kodierphase des Typenscoring hat sich der Rater nicht an strengen Zuordnungskriterien zu orientieren. Kohlberg betont den stärker klinischen Charakter des Verfahrens: "... the criteria are coded using a guided clinical evaluation. An individual response may not meet the exact definitions of a criterion for a particular dilemma, so the coder must be guided by the examples. The examples assist in the clinical evaluation of whether a response meets the definition of the criterion. ... Since coding for moral types involves the content and the structure of the reasoning rather than specified norms, elements, and structural definitions as is found in the Standard Issue Scoring Manual, a more clinical approach is taken" (Colby & Kohlberg, 1987; S. 911).

Auch die Kodierung des Interviewmaterials auf der Grundlage von Typenkriterien stellt relativ hohe Anforderungen an den Rater. Wie beim Standard Issue Scoring ist eine gute Kenntnis des Manuals und eine Vertrautheit mit den zugrundeliegenden theoretischen Konzepten notwendig.

Weil das Verfahren zur Kodierung von Typen ähnlich wie das frühe Standard Issue Scoring und das eben beschriebene Verfahren der Vergabe von Guess-Werten einen starken klinischen Charakter besitzt, sind die Unterschiede in den Einschätzungen von Ratern, die zu abweichenden Kodierergebnissen führen, hier nicht direkt zu greifen. Unterschiede in der Bewertung von Äußerungen werden nicht am Verhältnis der Rater zu operationalisierten Zuordnungskriterien deutlich.

In der ersten Kodierphase bei beiden Verfahren sind die 3 Kriterien für moralische Urteile anzuwenden. Die Differenzierung zwischen moralischen und nichtmoralischen Äußerungen ist nicht immer eindeutig vorzunehmen. Insbesondere ist es häufig schwierig, die Aussagen auf ihren präskriptiven Gehalt hin zu bestimmen und die subjektive Validität der Argumente einzuschätzen. Entsprechende Unterschiede in der Bewertung von Äußerungen können in beträchtlichem Ausmaß zu abweichenden Kodierergebnissen führen.

Die erste Kodierphase beim Standard Issue Scoring beinhaltet eine zusätzliche Abweichungsursache. Ein wesentlicher Bestandteil ist hier nämlich die Klassifikation des Interviewmaterials nach Issues. Der Rater hat die Interviewaussagen des Probanden zu einem Dilemma danach zu differenzieren, für welche der konkurrierenden Werte die Aussagen Gründe liefern.

Da die ersten beiden Kodierphasen bei jedem Verfahren hohe Anforderungen an die Rater stellen, ist zu erwarten, daß unerfahrene Rater bei der Kodierung von CJs bzw. Typenkriterien keine hohe Interraterreliabilität erzielen.

Im Hinblick auf das Standard Issue Scoring muß bezweifelt werden, daß es Kohlberg gelungen ist, durch die Formulierung von strengen Zuordnungskriterien auch relativ unerfahrenen Ratern eine reliable Kodierung zu ermöglichen. Der Versuch, den Auswertungsprozeß zu standardisieren, hat zu einer Vielzahl von Kodierregeln geführt, die Rater ohne große Kodierpraxis nicht vollständig berücksichtigen können. An die Stelle der Unsicherheiten des unerfahrenen Raters im Umgang mit dem stärker klinischen Structural Issue Scoring sind die Unsicherheiten des unerfahrenen Raters im Umgang mit dem psychometrisch durchstandardisierten Standard Issue Scoring getreten.

Wir haben in ersten Ansätzen die Frage zu klären versucht, wieweit beim Standard Issue Scoring erfahrene Rater eine hohe Interraterreliabilität bei der Kodierung von Criterion Judgments erzielen können. Wir wollten klären, ob die Interraterreliabilitäten für die CJ-Werte ähnlich hoch liegen wie die Interraterreliabilitäten, die Kohlberg und die an Kohlberg orientierte Forschung für die Form-Werte ermittelt haben. Darüber hinaus wollten wir prüfen, welche spezifischen Kodierschritte für abweichende Kodierergebnisse vor allem verantwortlich sind.

Eine Prüfung der Reliabilität, die unerfahrene Rater bei der Kodierung von CJs erzielen, wurde bislang noch nicht vorgenommen. Auch für die Kodierung der Typenkriterien wurden bislang keine Reliabilitätswerte ermittelt.

In unserer Untersuchung wurde die Reliabilität für die Kodierung von CJs zunächst von zwei Ratern geprüft, die mehrere Monate Kodierpraxis hatten und mit Kohlbergs theoretischen Konzepten vertraut waren. Vorrangiges Ziel der Rater war aber nicht die systematische Untersuchung der Interraterreliabilität. Ihre hauptsächliche Aufgabe bestand darin, eine breite Datenbasis für die Analyse von CJ-Werten zu schaffen.

Auf der Grundlage von Kohlbergs eigenen Daten war, wie schon erwähnt, eine differenzierte Analyse der CJ-Werte nicht möglich, weil uns diese Daten nur in defizitärer Form vorliegen. 20 Fälle wurden von den beiden Ratern rekodiert, wobei die Rekodierungen im Hinblick auf 4 Dilemmata (nämlich "Heinz", "Joe", "Euthanasie" und "Korea") erfolgten.

Das Ausmaß an Übereinstimmung, die diese erfahrenen Rater bei der Kodierung von CJs erzielten, war recht gering. Auch der Anteil der Übereinstimmung mit Kodierungen, die in den Auswertungsbögen von Kohlberg ausgewiesen sind, war nicht höher. Die Interraterreliabilität lag jeweils deutlich unter den für die Form-Werte ermittelten 80%.

Dieses Resultat veranlaßte uns, dem Problem der Interraterreliabilität intensiver nachzugehen. Die relativ geringe Übereinstimmung legte den Gedanken nahe, daß selbst bei erfahrenen Ratern die subjektive Willkür im Kodierprozeß recht groß ist. Es kamen Zweifel auf, ob die von uns angestrebte Mikroanalyse moralischen Urteilens überhaupt möglich ist.

Wir haben uns gefragt, wieweit die geringe Interraterreliabilität auf mögliche Fehler, die den Ratern in jeder Phase der Auswertung unterlaufen können, zurückzuführen ist. Dadurch daß (a) die Phase der Eintrainierung der Rater um einige Monate verlängert wurde, (b) die Rater sich nur auf bereits vorgenommene Kodierungen zu konzentrieren hatten und (c) die Kodierergebnisse (bei einigen Fällen) im Prozeß der Diskussion ermittelt wurden, sollte der Fehleranteil bei der Auswertung reduziert und damit die Qualität der Datenbasis gesteigert werden.

Drei neue Rater haben dann insgesamt 15 der von den ersten beiden Ratern rekodierten Fälle nochmals rekodiert, indem sie zu den vorgenommenen Kodierungen kritisch Stellung genommen haben. Sie haben jeweils festgehalten, bei welchen Kodierschritten sie von den früheren Kodierungen abweichen. Bei 3 Fällen wurde das Kodierresultat von den Ratern gemeinsam ermittelt. Bei den anderen Fällen hat sich jeder Rater einzeln mit den ursprünglichen Kodierungen auseinandergesetzt. Für 9 Fälle wurde das Ausmaß der Übereinstimmung mit den Kodierungen der früheren Rater untersucht. Bei 6 von 12 Fällen, die die Rater allein rekodiert haben, wurde im Hinblick auf das Heinz- und Joe-Dilemma die Übereinstimmung mit den ursprünglichen Kodierungen geprüft, wobei nicht nur die Reliabilität der Kodierung von CJs, sondern auch die Reliabilität der Kodierung von isolierten Normen, Elementen und Stufen ermittelt wurde. Für die 3 Fälle, bei denen die Rater das Kodierergebnis im Prozeß der Diskussion gewonnen haben, liegt gesondert eine Analyse des Anteils der möglichen Ursachen für abweichende Kodierergebnisse vor.

Der Prozentanteil an Übereinstimmung mit den früheren Ratern lag bei den 6 Fällen, die durch einen einzelnen Rater rekodiert wurden, zwischen 60% und 90%. Die für die Kodierung von CJs ermittelten Werte lagen dabei etwas niedriger als die für die Kodierung von isolierten Normen, Elementen und Stufen ermittelten Werte. Die Prozentanteile für die jeweiligen Normen, Elemente und Stufen unterscheiden sich kaum.

Bei den 3 Fällen, wo das Kodierergebnis von den Ratern im Prozeß der Diskussion gewonnen wurde, lag der Anteil der Übereinstimmung mit den ursprünglichen Kodierungen ebenfalls zwischen 60% und 90%. Die hier durchgeführte Analyse von Ursachen für abweichende Kodierungen von CJs zeigte, daß ein großer Teil der Abweichungen auf Unterschiede in der Erfassung moralischer Aussagen und in der Anwendung der critical indicators zurückzuführen ist.

Es war festzustellen, daß ein größerer Anteil von Abweichungen nicht eindeutig auf Auswertungsfehler zurückgeführt werden kann. Bruchstückhafte und mehrdeutige Interviewtexte sowie Ungereimtheiten und Unklarheiten im Manual sorgen häufig für kaum zu behebende Unsicherheiten bei der Kodierung.

Mit der Rekodierung von Fällen durch die drei intensiv eintrainierten Rater konnten Fehler bei der Auswertung getilgt und die Qualität der Datenbasis gesteigert werden. Andererseits aber wurde deutlich, daß eine Steigerung der Reliabilität bei erfahrenen Ratern nur begrenzt möglich ist.

#### **4. Zusammenfassung und Ausblick**

In der hier berichteten Rekonstruktion des Ansatzes und des Meßverfahrens von L. Kohlberg haben wir uns um weiterführende konzeptuelle Differenzierungen und Abklärungen im kognitiven Bereich der Moralentwicklung bemüht. Obwohl die Beschränkung auf den kognitiven Bereich im Rahmen einer Kohlbergrekonstruktion sicherlich berechtigt ist, bleiben deren Problemstellungen und Ergebnisse doch prinzipiell dadurch begrenzt, daß dabei von der Entwicklung moralischer Affekte und Motive und der konstitutiven Bedeutung, die diese Entwicklung für die Genese des moralischen Selbst hat, abstrahiert wird.

Die hier berichtete Kohlbergrekonstruktion erfolgte im Schnittpunkt von drei Analyseebenen: auf begriffs- bzw. bedeutungsanalytischer, auf normativ-analytischer und auf empirisch-analytischer Ebene. Nicht nur jede Kohlbergrekonstruktion, auch jede Weiterführung des Kohlbergschen Forschungsansatzes muß, sofern sie nicht von einer dogmatischen Festschreibung der Kohlbergschen Stufenbestimmungen und des Kohlbergschen Instrumentariums der Datengewinnung (moral judgment interview) und der Datenauswertung (standard issue scoring) ausgehen will, diese drei Analyseebenen berücksichtigen. Dies gilt freilich nur für eine Weiterführung, die nicht empiristisch verkürzt ist, sondern davon ausgeht,

daß der durch die Studie von Kramer (1968) in Gang gebrachte "bootstrapping-Prozeß" (s. Kap. 1.3) noch nicht abgeschlossen ist. Geht man von dieser Annahme aus, so besteht die vorrangige Forschungsaufgabe darin, konzeptuelle Fragen auf theoretischer wie operationaler Ebene abzuklären.

Im folgenden sollen in Beschränkung auf den kognitiven Bereich der Moralentwicklung Kapitel für Kapitel die entscheidenden Ergebnisse der hier durchgeführten Rekonstruktion referiert und mit dem Ausblick auf jeweils noch anstehende Aufgaben verknüpft werden. Im Anschluß daran soll eine Gesamtbewertung der Kohlbergschen Errungenschaften im Bereich kognitiver Moralentwicklung vorgenommen und schließlich eine über Kohlbergs kognitiv-strukturtheoretischen Forschungsansatz im engen Sinne hinausführende Forschungsperspektive umrissen werden.

Zu 2.1: *Vermittlung von rationaler und phänomenologischer Rekonstruktion.* In diesem Kapitel wurde primär, wenn auch nicht ausschließlich, unter normativ- und begriffsanalytischen Gesichtspunkten die Angemessenheit der theoretischen sowie der operationalen Stufenbestimmungen Kohlbergs problematisiert.

Auf präkonventionellem Niveau und hinsichtlich einer angemessenen Abgrenzung des postkonventionellen vom konventionellen Niveau moralischer Urteilsbildung erwiesen sich sowohl die theoretischen als auch die operationalen Stufenbestimmungen als problematisch; auf konventionellem Niveau hingegen zeigten sich nur auf der Ebene der operationalen Stufenbestimmungen gewisse Problembereiche. ("operational" definiert sich dabei lediglich durch die CJ-Ebene, genauer gesagt durch die CJs per se oder die CJs in Verknüpfung mit spezifischen "Stufenstrukturen").

Auf den Kohlbergstufen 1 und 2 des präkonventionellen Niveaus ist die Bestimmung der Stufenstrukturen inhaltlich verengt: auf Stufe 1 im Sinne der Erwartungen von Autoritätspersonen und auf Stufe 2 im egoistisch-instrumentalistischen Sinne. Bei der Bestimmung der Stufe 1 bleibt unklar, wie das Konzept des Egozentrismus zu verstehen ist: kann bei der Konfliktwahrnehmung bereits zwischen unterschiedlichen Interessenlagen differenziert werden, doch bei der Konfliktlösung finden nur die eigenen Interessen Berücksichtigung. Unklar bleibt auch, wie Personen mit dem Konflikt zwischen den (faktischen) Erwartungen und Regelsetzungen von Autoritätspersonen und elementaren, allgemeinen Normen umgehen und inwiefern Personen auf Stufe 1 schon in der Lage sind, zu unterscheiden zwischen (konventionellen) Regeln, die nur für bestimmte Personengruppen gelten und veränderbar sind und (moralischen) Normen, die für jedermann "absolut"



verbindlich sind und von niemandem, auch von Autoritätspersonen nicht, außer Kraft gesetzt, modifiziert oder eingeschränkt werden dürfen. Zudem läßt Kohlberg außer acht, daß die Verbindlichkeit moralischer Normen nicht selbstevident ist, sondern schon von kleinen Kindern unter empathischen Gesichtspunkten begründet werden kann (vgl. Turiel, 1983).

Die mangelnden kognitiv-strukturtheoretischen Differenzierungen bei der Bestimmung der Stufe 1 beruhen wahrscheinlich vor allem auf der Unterrepräsentation dieser Stufe in der Stichprobe Kohlbergs und folglich in seinem Auswertungsmanual.

Es stellt sich deshalb die Aufgabe, über unsere relativ globalen alternativen Stufenkonzepte hinaus auf der Ebene der theoretischen sowie der operationalen Stufenbestimmungen die erforderlichen konzeptuellen Differenzierungen vorzunehmen, wie sie durch die Untersuchungen von Damon (1984), Eisenberg (1979), Hoffmann (1984), Keller, Eckensberger & von Rosen (1989), Keller (1990), Nunner-Winkler & Sodian (1988) und Turiel (1983) nahegelegt werden.

Im Rahmen kulturvergleichender Forschung ist zu untersuchen, inwiefern sich die Begründungen für "absolut" verbindliche moralische Normen sowie die Differenzierungen, die zwischen veränderbaren konventionellen Regeln und "absoluten", unabänderbaren moralischen Normen vorgenommen werden, als kulturabhängig erweisen.

Wir haben bei den Stufen 1 und 2 des präkonventionellen Niveaus Bestimmungen der invarianten Tiefenstruktur vorgeschlagen, die Kohlbergs inhaltliche Verengungen - bei Stufe 1 auf Erwartungen von Autoritäten, bei Stufe 2 im egoistisch-instrumentalistischen Sinne - zu vermeiden suchen. Nur unter der Voraussetzung, daß Kohlbergs inhaltlich verengte Stufenbestimmungen auf präkonventionellem Niveau überwunden und durch angemessene Struktur-Inhaltsdifferenzierungen ersetzt werden, kann im Rahmen kulturvergleichender Forschung sinnvollerweise danach gefragt werden, inwieweit bereits auf präkonventionellem Niveau stufenabhängig und kulturabhängig variable inhaltliche Orientierungen (empathisch vs. egoistisch-instrumentell, beziehungs- oder gerechtigkeitsorientiert) auftreten, in denen sich universelle, invariante Tiefenstrukturen durchhalten (vgl. Keller, 1990; Keller & Edelstein, 1992).

Einen auf diese Fragen bzw. Untersuchungsgesichtspunkte zugeschnittenen Interviewleitfaden zu konstruieren, ist eine weitere Aufgabe. Es würde sich wahrscheinlich empfehlen, diesen Leitfaden so zu konstruieren, daß er den Probanden veranlaßt, eine Konfliktsituation nicht nur aus der Perspektive eines Aktors, sondern aus der Perspektive aller

in einen moralischen Konflikt involvierten Personen zu verstehen, sich die unterschiedlichen Handlungspräferenzen sowie die Folgen, die sie haben, zu vergegenwärtigen und zugleich eine moralische Bewertung der unterschiedlichen Handlungspräferenzen und ihrer Folgen vorzunehmen, die der Proband selbst für richtig und maßgeblich hält.

Außerdem sollten Kohlbergs Dilemmata (von "Joe" und "Judy" abgesehen) in eine kindgemäßere Form gebracht werden.

Auf dem konventionellen Niveau moralischer Urteilsbildung war lediglich die Operationalisierung der Stufenbestimmungen - vor allem hinsichtlich der Problemzonen "Bedeutungsunterstellung" und "Bedeutungsverschiebung" - ein Gegenstand der Kritik. Es wurden z.B. Bedeutungsunterstellungen nachgewiesen, die darauf beruhen, daß konkrete Rollenerwartungen, wie sie für Stufe 3 konstitutiv sind, in problematischer Vereindeutigung im Rahmen dekontextualisierter Auswertung von vornherein als Bestandteile einer Systemperspektive interpretiert und damit Stufe 4 zugewiesen worden sind.

Außerdem wurde die strukturelle Differenzierung zwischen "konkreter" und "abstrakter Systemperspektive" eingeführt. (Rollenerwartungen, die sich explizit hinsichtlich eines übergeordneten sozialen Ganzen definieren, sind i.S. einer konkreten Systemperspektive zu verstehen. Im Unterschied dazu wird bei einer abstrakten Systemperspektive von konkreten Rollenerwartungen abstrahiert und lediglich auf die allgemeinen Funktionsbedingungen sozialer Institutionen und Systeme Bezug genommen).

Mittels der Differenzierung zwischen konkreten Rollenerwartungen, die nicht in hierarchischer Integration Bestandteile einer Systemperspektive sind (Stufe 3) und konkreter und abstrakter Systemperspektive (Stufe 4) wurden beim Vergleich des personalen mit dem transpersonalen Dilemmatypus Bedeutungsverschiebungen nachgewiesen.

Kontextabhängige Bedeutungsverschiebungen legen aber die Schlußfolgerung nahe, sie seien als ein Indikator für das Scheitern von Kohlbergs Bemühungen zu verstehen, bereichsübergreifende, invariante Tiefenstrukturen zu ermitteln. Wie läßt sich eine derartige Schlußfolgerung entkräften und erfolgreich zurückweisen?

Zum einen dadurch, daß anhand hermeneutischer Einzelfallanalysen der Nachweis dafür erbracht wird, daß sich Bedeutungsverschiebungen identifizieren und im Rahmen einer Kontextualisierung der Bedeutung dadurch eliminieren lassen, daß problematische Intervieweinheiten nicht in atomistischer Auswertung unter die dafür im Manual

vorgesehenen CJs subsumiert, sondern an eine vorherrschende oder zumindest in hinreichend klarer, expliziter Form entwickelte Begründungsstruktur (z.B. der Stufe 3) assimiliert werden.

Zum anderen müßte ein Interviewleitfaden entwickelt werden, der eine größere Vergleichbarkeit zwischen unterschiedlichen Dilemmatypen (personal vs. transpersonal) schafft, indem er vergleichbare Situationsbedingungen herstellt. Bei personalen Dilemmata müßte der Proband veranlaßt werden, interpersonelle Beziehungskonflikte auch unter den Handlungsbedingungen sozialer Institutionen und im Hinblick auf die Funktionsbedingungen sozialer Systeme zu reflektieren, und zwar provoziert durch hartnäckig nachhakende Erläuterungsfragen und Begründungsfragen in möglichst ausdifferenzierter, elaborierter Form. Bei transpersonalen Dilemmata müßte im Gegensatz dazu, aber in komplementärer Form, zusätzlich zu den üblichen Kohlbergschen Standardfragen auch Fragen gestellt bzw. zusätzlich Handlungsbedingungen eingeführt werden, die den Probanden dazu veranlassen, sich auf die Relevanz interpersoneller Beziehungskonflikte einzulassen.

Wenn beim strukturellen Interview wie bei der Auswertung die vorgeschlagenen Wege besprochen werden, kann sinnvoll untersucht werden, inwieweit es sich bei den unterschiedlichen Perspektiven (Beziehungsperspektive, konkrete und abstrakte Systemperspektive), d.h. bei deren möglicher Konstruktion und Übernahme um kulturabhängige, alternative Entwicklungspfade handelt.

Auf dem postkonventionellen Niveau moralischer Urteile zeigte sich, daß sowohl Kohlbergs theoretische Abgrenzung von konventioneller und postkonventioneller Moral als auch deren Operationalisierung problematisch ist. Die Problematisierung erfolgte im Rahmen eines Klassifikationssystems, das auf der Unterscheidung zwischen "universalistischen Begründungen autonomer vs. nicht-autonomer Art" und "autonomen Begründungen universalistischer vs. nicht-universalistischer Art" beruht und in vollständiger Disjunktion alle denkbaren Fälle postkonventioneller Begründungsstrukturen zu erfassen und auszuschöpfen erlaubt. Auf postkonventioneller Ebene wurden über Kohlbergs Kennzeichnung der Stufe 4/5 hinausgehende, strukturelle Differenzierungen entwickelt, die es ermöglichen, für den Übergang von konventioneller zu postkonventioneller Moral alternative Entwicklungspfade zu spezifizieren.

Bestimmte Begründungsstrukturen, wie sie für die möglichen Übergänge von konventioneller zu postkonventioneller Moral charakteristisch sind, nämlich "universalistische Begründungen nicht-autonomer Art" sowie "universalistische Begründungen, die nur im schwachen Sinne autonomer Art sind", sind in Kohlbergs Auswertungsmanual unterrepräsentiert. Um diese

Begründungsstrukturen im Rahmen kulturvergleichender Forschung angemessen erfassen zu können, stellt sich deshalb zunächst einmal die Aufgabe, sie im Manual hinreichend zu repräsentieren.

Nicht nur Verbesserungen des Auswertungsmanuals, sondern auch des Instruments der Datengewinnung, d.h. der Dilemmakonstruktion und des Interviewleitfadens, sind jedoch für eine angemessene Erfassbarkeit der genannten Begründungsstrukturen erforderlich: Zum einen wäre eine Konstruktion bzw. Vorgabe von Dilemmata sinnvoll, in der Norm- und Interessenkonflikte aus der Kollision von inkompatiblen Formen partikularer Gruppenmoral entstehen, zum anderen müßten die Interviewfragen auf das Verständnis dafür zielen, ob und inwiefern die genannten Begründungsstrukturen bei derartigen Dilemmata zur (moralischen) Konfliktlösung bzw. zur Konfliktvermeidung beitragen oder ihrerseits noch eine Quelle von Kollisionen sind - Kollisionen, die nur unter der Voraussetzung erfolgreich vermieden oder beseitigt werden können, daß es zur Ausdifferenzierung von unterschiedlichen Wert- und Geltungssphären (Religion, Recht, Moral, Konvention usf.) in ihrer Eigensinnigkeit und zum Rekurs auf universelle Personen- und Freiheitsrechte kommt (Stufe 5).

Im Spannungsfeld von rationaler und phänomenologischer Rekonstruktion wurde - um der phänomenologischen Rekonstruktion ein stärkeres Gewicht zu verschaffen, als dies bei Kohlberg der Fall ist - eine Weiterentwicklung des Restschen Präferenztests (DIT, 1983) vorgenommen.

Es wäre eine sinnvolle Aufgabe, diese Weiterentwicklung dadurch hinreichend zu konkretisieren, daß sie zum integralen Bestandteil eines komplexeren Interviewleitfadens bei der Konstruktion struktureller Interviews gemacht würde - eine Konstruktion, die sich für die Annahme von kulturabhängigen, alternativen Entwicklungspfaden offen hält.

Zu 2.2: *Vermittlung von hermeneutischer und psychometrischer Forschungsstrategie.* In diesem Kapitel erfolgte vor allem eine Kritik an Kohlbergs atomistischem Auswertungsverfahren (standard issue scoring) unter hermeneutischen Gesichtspunkten. Zugleich wurde ein Interpretationsansatz für Stufeninkonsistenzen entwickelt, der diese i.S. kohärenter Sinn Ganzheiten zu verstehen erlaubt.

Die Grenzen eines atomistischen Auswertungsverfahrens und der präjudikativen Funktion, die Kohlbergs Auswertungsmanual dabei erfüllen soll, werden zumindest hinsichtlich der Problemzonen "Bedeutungsunterstellung", "Bedeutungsverschiebung" und "mangelnde Differenzierung zwischen dem kognitiven Entwicklungsniveau und der sprachlichen Elaboration

moralischer Urteile" manifest. Zumindest bei diesen Problemzonen wird für die strukturelle Evaluierung eine Kontextualisierung der Bedeutung erforderlich.

Im Rahmen hermeneutischer Einzelfallanalysen tritt an die Stelle isolierter Argumentationsstufen der Versuch, unterschiedliche Argumentationsstufen dadurch sinnvoll aufeinander zu beziehen, daß zu klären versucht wird, welche Funktion und Bedeutung einzelne Argumentationsschritte in komplexen Argumentationsgängen haben. Die Rekonstruktion dieser Funktion und Bedeutung erfolgt im Lichte des Interpretationsgesichtspunkts situationsadäquater Urteilsbildung.

Zum einen wäre es wichtig, die hinsichtlich der Guess-scores und der Verrechnungsregeln an Kohlbergs Auswertungsverfahren geübte Kritik anhand von Interviewbeispielen zu illustrieren und zu vertiefen.

Zum anderen sollte anhand des Kohlberg-Manuals und des Gebrauchs, der im Rahmen atomistischer Auswertung davon zu machen ist, sowie anhand eines revidierten Manuals bzw. mittels einer Manualverwendung, die der erforderlichen Kontextuierung der Bedeutung Rechnung trägt, unter differentiellen (typologischen) Gesichtspunkten untersucht werden, ob es sinnvoll ist, Personengruppen durch den Gebrauch zu unterscheiden, die sie von bestimmten Argumentationsstufen machen. Inwieweit lassen sich Personengruppen identifizieren, die sich dadurch unterscheiden, daß sie Entscheidungen in relativ flexibler, situationsadäquater Form begründen?

Die von Walker & Taylor (1991) in ihrem Konsolidierungs-/Transitionsmodell unter globalen entwicklungsdynamischen Gesichtspunkten entwickelten Parameter können auch dazu verwendet werden, zwischen situationsadäquater und situationsinadäquater Urteilsbildung zu unterscheiden. Der psychometrische Untersuchungsansatz erlaubt aber lediglich die Erfassung von relativ groben Indikatoren situationsadäquater vs. situationsinadäquater Urteilsbildung. Um der erforderlichen Kontextuierung der Bedeutung Rechnung tragen zu können, ist eine Vertiefung und Illustration durch hermeneutische Einzelfallanalysen erforderlich, die auf eine Rekonstruktion typenspezifischer Argumentationsmuster zielen.

Zu 2.3: *Analyse des Standard Issue Scoring-Manuals*. Bereits in den vorangegangenen Kapiteln (2.1. und 2.2) wurden wichtige Ergebnisse der Manualanalyse vorweggenommen: unter strukturellen Aspekten werden die 3 Problembereiche "Bedeutungsunterstellung", "Bedeutungsverschiebung" und "mangelnde Differenzierung zwischen dem kognitiven Entwicklungsniveau und der sprachlichen Elaboration moralischer Urteile" aufgezeigt. Es

wurde hervorgehoben, daß zumindest hinsichtlich dieser 3 Problembereiche Kohlbergs Auswertungsmanual nicht die im Rahmen eines atomistischen Auswertungsverfahrens nötige präjudikative Funktion, sondern lediglich eine heuristische Funktion erfüllen kann und der Evaluierungsvorgang eine Kontextuierung der Bedeutung erfordert.

Hinsichtlich der Inhaltsaspekte moralischer Entscheidungsbegründungen wurde bereits in Kapitel 2.1 darauf hingewiesen, daß alle 4 Elementegruppen Kohlbergs (1. "Aufrechterhaltung der normativen Ordnung", 2. "utilitaristische Folgenorientierung", 3. "perfektionistische Orientierung" und 3. "Fairneßorientierung") lediglich auf dem konventionellen Niveau moralischer Urteilsbildung auftreten, wohingegen auf dem präkonventionellen Niveau nur die ersten beiden Elementegruppen und auf dem postkonventionellen Niveau (Stufe 4/5 und 5) ausschließlich die 3. und 4. Elementegruppe vorhanden sind.

Das hat aber in Blick auf die hierarchische Integration der Stufenbestimmungen erhebliche Einschränkungen der Kohlbergschen Stufenkonstruktion zur Folge. Um diese Defizite zu beseitigen, wurden deshalb neue, stufenunabhängige Inhaltsdimensionen vorgeschlagen.

Diese Ergebnisse der Manualanalyse mußten für die Klärung der Probleme, die sich in den Kapiteln 2.1. und 2.2. stellen, vorweggenommen werden.

In den Kapiteln 2.1. und 2.2. wurden lediglich das Verhältnis von CJs zu den Stufenstrukturen (Stage Structures) sowie Inklusionen berücksichtigt. Außerdem wurde nur auf die reinen Stufen Bezug genommen, ausgenommen die Übergangsstufe 4/5. In dem Kapitel 2.3 hingegen sind auch die kritischen Indikatoren und prototypischen Beispiele (match examples) sowie die Zwischenstufen als Bestandteile der Manualkonstruktion analysiert worden. Außerdem wurde in 2.3 untersucht, inwieweit eine (mehr oder weniger stark ausgeprägte) Ungleichverteilung der CJs in Abhängigkeit (a) von Niveaus und Stufen, (b) von Fragetypen (situationsspezifische vs. konzeptuelle) und (c) von Issues (autonom vs. heteronom) auftritt.

Es zeigte sich, daß im Vergleich zu den CJs auf konventionellem Niveau vor allem präkonventionelle CJs weniger ausgeprägt, aber auch postkonventionelle CJs unterrepräsentiert sind.

Auf präkonventionellem Niveau ist zugleich das autonome Issue unterrepräsentiert, während auf postkonventionellem Niveau das heteronome Issue unterrepräsentiert ist.

In Beziehung auf die konzeptuellen Fragen konnte festgestellt werden, daß bei einigen Dilemmata und einigen Issues keine oder fast keine konzeptuellen Fragen auftreten - obwohl im Fragenkatalog des standardisierten Interviews konzeptuelle Fragen enthalten sind, z.B. beim Heinz-Dilemma, Issue: "Leben", Euthanasie-Dilemma, Issue: "Lebensqualität".

Zwar erwies sich die Passung von prototypischen Beispielen (match examples) und CJs, nicht aber das Verhältnis von kritischen Indikatoren und Stufenstrukturen (stage structures) als unproblematisch. Inwiefern die Stufenstrukturen die ihnen zugeordnete Funktion "struktureller Evaluation" in Kohlbergs Auswertungsverfahren erfüllen und im Vergleich zu den kritischen Indikatoren präzisere und strukturtheoretisch ausgewiesene Stufenzuordnungen ermöglichen sollen, ist fragwürdig. Entweder reduzieren sich nämlich die Stufenstrukturen auf wörtliche Wiederholungen der bereits bei den kritischen Indikatoren erfolgten Kennzeichnungen oder auf allgemeinere Formulierungen, oder sie stellen - sofern sie tatsächlich neue, zusätzliche Bestimmungen ins Spiel bringen - theoretische Überinterpretationen bzw. Bedeutungsunterstellungen dar.

In hohem Maße problematisch in Kohlbergs Manualkonstruktion sind auch die Strukturbestimmungen der Zwischenstufen: diese sind nur in wenigen Fällen (CJs) positiv als qualitativ eigenständige Begründungsformen bestimmt, relativ häufig hingegen werden sie als Argumentationsstrukturen bestimmt, die nicht mehr für eine Stufe (n), aber auch noch nicht für eine Stufe (n+1) kennzeichnend sind, oder aber ihre Unterbestimmtheit bzw. mangelnde Bestimmbarkeit wird auf Mängel der Interviewführung zurückgeführt.

Folgende Aufgaben stellen sich: anstelle der bisher z.T. nur exemplarisch durchgeführten Manualkritik muß eine alle 3 Formen umfassende, systematische Manualkritik bzw. eine systematische Überprüfung der evaluativen Funktion des Kohlbergschen Filtersystems ("Stufenstrukturen", "kritische Indikatoren" und "Inklusionen") treten, freilich nur, sofern eine solche Überprüfung in den Grenzen einer Manualanalyse möglich ist. Außerdem ist in entsprechender systematischer Form im Rahmen einer Manualanalyse zu untersuchen, wie die Zwischenstufen gekennzeichnet werden.

Über die Grenzen einer Manualanalyse hinausgehend sollten die genannten Problembereiche durch Beispiele, die anhand hermeneutischer Einzelfallanalysen exemplarisch gewonnen worden sind, illustriert werden, wobei zugleich die Frage zu klären ist, inwiefern eine Kontextuierung der Bedeutung eine überzeugende, konstruktive Bewältigung dieser Problemzonen erlaubt: sei es durch die Assimilation von unterbestimmten oder nur per negationem

bestimmten Zwischenstufen an reine Stufen oder durch die Eliminierung von Bedeutungsunterstellungen.

*Zu 2.4: Das mehrdimensionale Modell.* Das in diesem Kapitel vorgestellte "mehrdimensionale Modell" reagiert auf gewisse, "progressive Problemverschiebungen" (Lakatos), die in den letzten Jahren im Forschungsprogramm kognitiv-strukturtheoretisch orientierter Moralpsychologie erfolgt sind, indem es über Kohlbergs deontisch-präskriptive Verengung des sozialkognitiven Bereichs der Moralentwicklung, ja über den kognitiven Bereich überhaupt hinausführt.

Die in den letzten Jahren erfolgte Fokussierung auf die Entwicklung des moralischen Selbst (s. Edelstein & Nunner-Winkler, 1993) führte dazu, daß Kohlbergs einseitige Betonung der Entwicklung moralischer Urteilsfähigkeit und die Reduktion moralischer Urteile auf die rein präskriptiv-deontische Dimension von Gerechtigkeitsurteilen der Tendenz nach korrigiert wurde, wenn auch nur in Ansätzen.

Das mehrdimensionale Modell der moralischen Handlungsfähigkeit geht von folgenden Überlegungen aus:

Für die Entwicklung moralischer Handlungsfähigkeit, durch die sich ein verantwortungsbeußtes moralisches Selbst auszeichnet, ist die Entwicklung moralischer Sensibilität sowie die Entwicklung der Fähigkeit und Bereitschaft, Verantwortung zu übernehmen, von konstitutiver Bedeutung. Eine angemessene Rekonstruktion der Entwicklung moralischer Handlungsfähigkeit kann nur in einem handlungstheoretischen Rahmen erfolgen. In diesem Rahmen muß der Vermittlung moralischer Handlungsfähigkeit durch moralische Urteilskraft in einem unverkürzten Sinne Rechnung getragen werden. Die Rekonstruktion der Entwicklung moralischer Handlungsfähigkeit verlangt deshalb, daß die Differenzierungen und Verknüpfungen, die zwischen deskriptiven und präskriptiven Urteilen, Geltungs- und Anwendungsfragen, deontischen Urteilen und Verantwortlichkeitsurteilen sowie zwischen kommunikativen (verständigungsorientierten) und strategisch-erfolgsorientierten Formen der Handlungskoordination hergestellt werden, gebührende Berücksichtigung finden.

Welche Aufgaben stellen sich im Rahmen eines multidimensionalen Modells? Zunächst einmal sind anhand der uns zur Verfügung stehenden Kohlbergdaten die einzelnen Dimensionen ("deskriptive" vs. "präskriptive Urteile", "Geltungs-" vs. "Anwendungsfragen",



"deontische Urteile" vs. "Verantwortlichkeitsurteile", "verständigungsorientierte" vs. "strategisch-erfolgsorientierte" Formen der Handlungskoordination) zu illustrieren - eine Aufgabe, die aufgrund der präskriptiv-deontisch verengten Problemstellung Kohlbergs freilich auf verhältnismäßig enge Grenzen stößt.

Zum anderen muß versucht werden, durch Berücksichtigung der einschlägigen Forschungsergebnisse anderer Autoren (z.B. Keller & Edelstein, 1986; Nunner-Winkler & Sodian, 1988; Selman, 1984) diese Grenzen aufzuheben.

Zu 3.1 bis 3.3: *Ergebnisse auf empirisch-analytischer Ebene.* Auf empirisch-analytischer Ebene erfolgte eine generelle Kritik daran, daß Kohlberg ebenso wie die an Kohlberg orientierte Forschung sowohl bei der Analyse der Stufen als auch bei der Analyse der Typen die niedrig aggregierten Werte vernachlässigt haben und auch die Inhaltsaspekte, d.h. die Normen, Elemente und Handlungspräferenzen, kaum Gegenstand der Analyse waren (s. 3.1).

Erste Ergebnisse zur Frage der Konsistenz und Kontextbedingtheit der Stufen auf dem Aggregationsniveau von Issue-Werten liegen vor. Es zeigte sich, daß es zu beträchtlichen Inkonsistenzen kommt, die Kohlbergs "Dominant Stage Model" in Frage stellen und daß ein Teil der auftretenden Inkonsistenzen sich durch den Einfluß von Issues und Dilemmata erklären läßt (Teo, Becker & Edelstein, 1995).

Schließlich wurde das Problem der Interraterreliabilität bei der Kodierung von CJs behandelt. Es wurden verschiedene Fehlerquellen bestimmt (z.B. "kritische Indikatoren nicht oder nicht ausreichend erfüllt"; "Stufenstrukturen nicht erfüllt", "Inklusionen nicht berücksichtigt") und das relative Gewicht der Abweichungsursachen ermittelt. Es zeigte sich, daß selbst bei erfahrenen Ratern ein hohe Interraterreliabilität auf CJ-Ebene nur schwer herstellbar ist und bei bestimmten Interviews selbst der Versuch einer gemeinsamen Abklärung von Kodierinhalten an seine Grenzen stößt, vor allem im Falle von Interviews, bei denen nicht hinreichend konsequent im Sinne eines strukturellen Interviews bohrende "Warum"-Fragen und Erläuterungsfragen verwendet wurden (s. 3.3).

Folgende Aufgaben stellen sich in Zukunft: auf der Ebene der Issue-Werte ist eine Vertiefung der Analysen hinsichtlich der Fragen der Konsistenz und der Kontextabhängigkeit erforderlich. Insbesondere muß eine differenziertere Beschreibung von Mustern der Inkonsistenz erfolgen.

Auf der Ebene der CJ-Werte ist eine systematische Analyse hinsichtlich der Fragen der Konsistenz, der Konsolidierung und der Kontextbedingtheit vorzunehmen. Auf inhaltlicher Ebene muß unter den Gesichtspunkten der Konsistenz und Kontextbedingtheit eine Analyse der Elemente und Handlungspräferenzen erbracht werden. Die schon weit vorangetriebene Mikroanalyse der Kohlbergschen Typenwerte muß anhand der vorliegenden Auswertungsdaten zu Form A zum Abschluß gebracht werden. Für die auf CJ-Ebene vorgesehenen empirischen Analysen sind die von uns rekodierten Fälle heranzuziehen.

Eine kritische Würdigung Kohlbergs, die den Leistungen dieses Forschers auf dem Gebiet der Moralphyschologie gerecht werden will, muß Kohlbergs Forschung primär an ihren eigenen, konstitutiven Ansprüchen messen.

Diese Ansprüche definieren sich dadurch, daß Kohlberg - analog zur kritischen Selbstbegrenzung der Vernunft bei Kant und in Fortführung der Kantschen Tradition - eine Selbstbegrenzung rationaler Rekonstruktion auf den präskriptiv-deontischen Bereich moralischer Urteilsfähigkeit, d.h. auf "Gerechtigkeitsurteile" vornimmt.

Kohlberg geht bei der letztgültigen, methodologischen Ortsbestimmung seines kognitiv-strukturtheoretischen Forschungsansatzes in den 80er Jahren davon aus, daß der für diesen Forschungsansatz konstitutive Anspruch rationaler Rekonstruktion sinnvollerweise nur erhoben und eingelöst werden kann, wenn zwei Voraussetzungen erfüllt sind: Zum einen, daß für die Rekonstruktion der Entwicklung moralischer Urteilsfähigkeit ein ausgezeichneter normativer Bezugspunkt, d.i. der Standpunkt einer autonomen und universalistischen Moral in der Tradition rationaler Normenbegründung unverzichtbar ist, zum anderen, daß damit eine doppelte Selbstbegrenzung kognitiv-strukturtheoretisch ausgerichteter Moralphyschologie impliziert ist - (a) eine Begrenzung auf den kognitiven Bereich der Moralentwicklung und (b) innerhalb dieses Bereichs eine Begrenzung auf die präskriptiv-deontische Dimension moralischer Urteile.

Unseres Erachtens handelt es sich zwar beim normativen Bezugspunkt einer autonomen (religionsunabhängigen und metaphysikfreien), universalistischen Moral in der Tradition rationaler Normenbegründung um eine notwendige Voraussetzung rationaler Rekonstruktion; die genannte rigoristische Selbstbeschränkung kognitiv-strukturtheoretisch ausgerichteter Moralphyschologie ist damit aber nicht notwendig impliziert. Der Versuch, Kohlbergs präskriptiv-deontisch verengtes Stufenmodell moralischer Urteilsfähigkeit in ein mehrdimensionales, handlungstheoretisches Modell der Entwicklung des moralischen Selbst zu integrieren, geht von dieser Einsicht aus.

Die kritischen Ergebnisse der Kohlbergrekonstruktion auf begriffs- bzw. bedeutungsanalytischer, auf normativ-analytischer und auf empirisch-analytischer Ebene führen zu der Schlußfolgerung, daß an Kohlbergs Forschungsansatz als progressivem Paradigma nur unter der Voraussetzung festgehalten werden kann, daß man mit Kohlberg über Kohlberg hinauszugehen versucht.

Weder ist eine Preisgabe der grundbegrifflichen, analytischen Unterscheidung zwischen den "Strukturen" und Inhalten" moralischer Urteile erforderlich noch die Aufgabe des darauf errichteten Stufenmodells und seiner Operationalisierung im "Standard Issue Scoring".

Das vor der Struktur-Inhaltsdistinktion bereits entwickelte Instrument der Datengewinnung (moral judgment interview) muß allerdings klarer und differenzierter auf eine angemessene Erfassung der Vielfalt variabler Inhaltsaspekte vs. invarianter Tiefenstrukturen zugeschnitten werden. Erst dann verdient das "strukturelle Interview" seinen Namen zurecht.

An den Errungenschaften, die die kognitiv-strukturtheoretisch orientierte Moralpsychologie Kohlberg zu verdanken hat, kann und sollte festgehalten werden. Es sind freilich Revisionen und Weiterentwicklungen erforderlich, zum Teil auf der theoretisch-konzeptuellen Ebene, zum Teil auf der operationalen Ebene.

Mit der progressiven Problemverschiebung, die in den letzten Jahren mit der Fokussierung auf die Entwicklung des moralischen Selbst (s. Edelstein, Nunner-Winkler & Noam, 1993) erfolgt ist, wird Kohlbergs einseitige Betonung der Entwicklung moralischer Urteilsfähigkeit und die Reduktion moralischer Urteile auf die rein präskriptiv-deontische Dimension von Gerechtigkeitsurteilen korrigiert, prinzipiell jedenfalls. Für die Entwicklung des moralischen Selbst ist die Entwicklung moralischer Sensibilität und die Entwicklung moralischer Handlungsfähigkeit, d.h. die Bereitschaft, Verantwortung zu übernehmen und Formen primär verständigungsorientierter Handlungskoordination zu entwickeln, von konstitutiver Bedeutung. In dem in seinen Grundlinien entworfenen "mehrdimensionalen Modell" wird deshalb versucht, der Vermittlung moralischer Handlungsfähigkeit durch moralische Urteilskraft in einem unverkürzten Sinne Rechnung zu tragen.

Eine angemessene Rekonstruktion der Entwicklung des moralischen Selbst macht die theoretische Klärung folgender Fragen notwendig:

- Welche Kultivierung moralischer Urteilskraft, und

- welche Kultivierung der moralischen Affekte hat eine fortschreitende Entwicklung moralischer Verantwortungs- bzw. Handlungsfähigkeit zur Voraussetzung, und
- welche makro- sowie mikrosozialen Entwicklungsbedingungen erweisen sich für die Schaffung dieser Voraussetzungen als mehr oder weniger förderlich bzw. hemmend?

Letztere Frage stellt sich von einem "sozialkonstruktivistischen" Forschungsansatz aus, wie er von Edelstein (1992) und Lempert (1988) entwickelt worden ist.

Es ist eine Forschungsfrage von hoher, auch praktischer Relevanz, die Fragen, die durch das Problem der Kultivierung der Affekte aufgeworfen werden, in Rückbesinnung auf die Traditionen der Moralphilosophie theoretisch und empirisch abzuklären. Dazu gehören in den vorkantischen Traditionen vor allem D. Hume, *Enquiries concerning the Human Understanding and concerning the Principles of Morals* und A. Smith, *The Theory of Moral Sentiments*; bei Kant selber vor allem *Die Metaphysik der Sitten* und in Beziehung auf nachkantianische Positionen z.B. E. Fromm *Die Furcht vor der Freiheit*, A. Schopenhauer, *Preisschrift über die Grundlage der Moral*, G. Taylor, *Pride, Shame and Guilt*; E. Tugendhat, *Vorlesungen über Ethik*.

Die (strukturell-internen) Bedingungen der Aufrechterhaltung bzw. Wiederherstellung der Integrität des moralischen Selbst, des eigenen wie des fremden, für die wir nicht nur als autonome Subjekte jeweils individuell, sondern gemeinsam verantwortlich sind, sind identisch mit den Bedingungen für die Aufrechterhaltung und Wiederherstellung der sozialen Gemeinschaft bzw. mit den Bedingungen für eine mögliche Fortsetzung moralischer Interaktion zwischen Subjekten, die sich als "vollwertige" Mitglieder der sozialen Gemeinschaft verstehen und behandeln, d.h. die sich gegenseitig achten können.

Das Verständnis dieser komplexen Struktur interner Bedingungen ist einerseits an

(a) negative moralische Affekte gebunden, d.h. an die Erfahrungen von Scham und Schuld, die an die Formen einer partiellen oder totalen Selbstverneinung rückgebunden sind. Sie kommen in Formen der Selbstverachtung, des Selbsthasses, der Selbstbestrafung, der kommunikativen Selbstisolierung wie der Selbstzerstörung sowie in entsprechenden destruktiven Affekten gegenüber anderen, z.B. des Fremdenhasses, der Aussonderung aus der sozialen Gemeinschaft oder in Zerstörungswut zum Ausdruck.

(b) Andererseits müssen dabei positive moralische Affekte sowohl vorausgesetzt als auch freigesetzt werden: Die Fähigkeit und Bereitschaft, den Anderen als Gleichberechtigten zu achten und Akte der tätigen Reue im Falle einer Verletzung der Integrität des Anderen zu vollziehen, komplementäre Akte der Verzeihung, der Vergebung und Versöhnung, aber auch Trauer und

Scham angesichts von Unrecht, das nicht vergeben werden kann und an dem die Akteure aufgrund der Zugehörigkeit zu einer historischen Schicksalsgemeinschaft teilhaben. Ohne die Freisetzung dieser Affekte und ohne das betroffene Wissen um sie als Voraussetzungen für eine mögliche Wiederherstellung moralischer Integrität und Interaktionsfähigkeit kann von der Entwicklung eines moralischen Selbst keine Rede sein.

## Literaturverzeichnis

- Apel, K. O. (1988). Diskurs und Verantwortung. Frankfurt a. M.: Suhrkamp.
- Ben-Chorin, Schalom (1977). Bruder Jesus. Der Nazarener aus jüdischer Sicht. München: dtv.
- Berkowitz, M. & Keller, M. (1994). Transitional processes in social cognitive development: A longitudinal study. International Journal of Behavioral Development , 17, 447-467.
- Blasi, A. (1983). Moral cognition and moral action. A theoretical perspective. Developmental Review, 3, 178-210.
- Boyes, M. & Walker, L. J. (1988). Implications of the cultural diversity for the universality claims of Kohlberg's theory of moral reasoning. Human Development, 31, 44-59.
- Colby, A. & Kohlberg, L. (1987). The measurement of moral judgment. Bd. 1: Theoretical foundations and research validation. New York: Cambridge University Press.
- Colby, A., Kohlberg, L., Speicher, B., Hewer, A, Candee, D. Gibbs, J. & Power, C. (1987). The measurement of moral judgment. Vol 2: Standard issue scoring manual. New York: Cambridge University Press.
- Colby, A., Kohlberg, L. & Gibbs, J. & Liebermann, M. (1983). A longitudinal study of moral judgment. Monographs of the Society for Research in Child Development, 48, 1-124.
- Damon, W. (1984). Die soziale Welt des Kindes. Frankfurt a. M.: Suhrkamp.
- De Vries, B. & Walker, L.J. (1988). Moral reasoning and attitudes toward capital punishment. Developmental Psychology, 22, 509-513.
- Döbert, R. (1987). Horizonte der an Kohlberg orientierten Moralforschung. Zeitschrift für Pädagogik, 33, 491-512.
- Döbert, R. & Nunner-Winkler, G. (1978). Performanzbestimmende Aspekte moralischen Bewußtseins. In G. Portele (Hrsg.), Sozialisation und Moral (S. 101-121). Weinheim: Beltz.
- Eckensberger, L. H. (1986). Handlung, Konflikt, Reflexion: Zur Dialektik von Struktur und Inhalt im moralischen Urteil. In W. Edelstein & G. Nunner-Winkler (Hrsg.), Zur Bestimmung der Moral (S. 409-442). Frankfurt a.M.: Suhrkamp.
- Eckensberger, L. H. (1989). Consistencies in inconsistencies. Poster presented at the session "Moral development", 10th biennial meeting of the International society for the Study of Behavioural Development (ISSBD), 9-13.7. 1989, Jyväskylä, Finland.
- Eckensberger, L. H. (1995). Moralische Urteile im Kulturvergleich. In: Thomas, A. (Hrsg.), Kulturvergleichende Psychologie. Göttingen: Hogrefe.
- Eckensberger, L. H. & Burgard, P. (1986). Zur Beziehung zwischen Struktur und Inhalt in der Entwicklung des moralischen Urteils aus handlungstheoretischer Sicht. Arbeiten der Fachrichtung Psychologie, 77. Saarbrücken: Universität des Saarlandes.
- Eckensberger, L. H. & Reinshagen, H. (1980). Kohlbergs Stufentheorie moralischen Urteilens: Ein Versuch ihrer Reinterpretation im Bezugsrahmen handlungstheoretischer Kon-

- zepte. In L. H. Eckensberger, R. K. Silbereisen (Hrsg.), Entwicklung sozialer Kognitionen: Modelle, Theorien, Methoden, Anwendung (S. 65-131). Stuttgart: Klett-Cotta.
- Edelstein, W. (1986). Moralische Intervention in der Schule. Skeptische Überlegungen. In F. Oser, R. Fatke, O. Höffe (Hrsg.), Transformation und Entwicklung. Grundlagen der Moralerziehung (S. 327-349). Frankfurt a. M.: Suhrkamp.
- Edelstein, W. (1992). Strukturgenese. Interne und externe Bedingungen der Entwicklung. Ein Beitrag zur Verbindung von sozialisationstheoretischen und entwicklungstheoretischen Fragestellungen. In: Beiträge des Forschungsbereichs Entwicklung und Sozialisation am Max-Planck-Institut für Bildungsforschung, Nr. 40, Berlin.
- Edelstein, W. & Hoppe-Graff, S. (1993): Die Konstruktion kognitiver Strukturen. Perspektiven einer konstruktivistischen Entwicklungspsychologie. Bern: Huber.
- Edelstein, W., Keller, M. & Schröder, E. (1990): Child development and social structure: A longitudinal study of individual differences. In P. B. Baltes, D. L. Featherman, R. M. Lerner (Hrsg.), Life-span development and behavior, Bd. 10, (S. 151-185).
- Edelstein, W. & Noam, G. (1982). Regulatory Structures of the Self and "Postformal" Stages in Adulthood. Human Development, 25, 407-422.
- Edelstein, W. & Nunner-Winkler, G. (1986). Zur Bestimmung der Moral. Frankfurt a. M.: Suhrkamp.
- Edelstein, W., Nunner-Winkler, G. & Noam, G. (1993). Moral und Person. Frankfurt a. M.: Suhrkamp.
- Eisenberg, N. (1979). The development of children's prosocial moral judgment. Developmental Psychology, 15, 128-137.
- Gilligan, C. (1984). Die andere Stimme. München: Piper.
- Habermas, J. (1983). Moralbewußtsein und kommunikatives Handeln. Frankfurt a.M.: Suhrkamp.
- Habermas, J. (1991). Erläuterungen zur Diskursethik. Frankfurt a.M.: Suhrkamp.
- Hart, D. (1992). Becoming Men. The development of aspirations, values and adaptional styles. New York, London: Plenum Press.
- Hoffman, M. (1984). Empathy, its limitations, and its role in a comprehensive moral theory. In W. Kurtines & J. Gewirtz (Hrsg.), Morality, moral behavior and moral development (S. 282-302). New York: Wiley.
- Keller, M. (1976). Kognitive Entwicklung und soziale Kompetenz: Zur Entwicklung der Rollenübernahme in der Familie. Stuttgart: Klett.
- Keller, M. (1984). Rechtfertigungen. Zur Entwicklung praktischer Erklärung. In W. Edelstein & J. Habermas (Hrsg.), Soziale Interaktion und soziales Verstehen (S. 253-299). Frankfurt a.M.: Suhrkamp.
- Keller, M. (1990). Zur Entwicklung moralischer Reflexion: Eine Kritik und Rekonzeptualisierung der Stufen des präkonventionellen moralischen Urteils in der Theorie von L. Kohlberg. In M. Knopf & W. Schneider (Hrsg.), Entwicklung Allgemeine Verläufe - Individu-

elle Unterschiede - Pädagogische Konsequenzen. [Festschrift zum 60. Geburtstag von Franz Emanuel Weinert] (S. 19-44). Göttingen: Hogrefe Verlag für Psychologie.

Keller, M. (1994). Moral und Beziehung: Die Entwicklung sozialer und moralischer Sensibilität von der Kindheit bis zum Jugendalter. Habilitationsschrift am Max-Planck-Institut für Bildungsforschung, Berlin.

Keller, M. (im Druck). Gerechtigkeit und Fürsorge: Die Entwicklung der Moral in Freundschaft und Familie. Psychologische Verlagsanstalt, Weinheim.

Keller, M., Eckensberger, L. H. & von Rosen, K. (1989). A critical note on the conception of pre-conventional morality: The case of stage 2 in Kohlberg's theory. International Journal of Behavioral Development, 12, 57-69.

Keller, M. & Edelstein, W. (1986). Beziehungsverständnis und moralische Reflexion. In W. Edelstein & G. Nunner-Winkler (Hrsg.), Zur Bestimmung der Moral (S. 321-346). Frankfurt a.M.: Suhrkamp.

Keller, M. & Edelstein, W. (1991). The development of socio-moral meaning making: Domains, categories, and perspective taking. In W. Kurtines & J. Gewirtz (Hrsg.), Handbook of moral behavior and development, Bd. 2 (S. 89-114). Hillsdale: Erlbaum.

Keller, M. & Edelstein, W. (1992). Zur Universalität moralischen Urteilens: Eine kulturvergleichende Untersuchung zum präkonventionellen moralischen Urteil. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 1, 3-21.

Keller, M. & Reuss, S. (1984). An action-theoretical reconstruction of the development of social cognitive competence. Human Development, 27, 211-220.

Keller, M. & Reuss, S. (1986). Der Prozeß moralischer Entscheidungsfindung. Normative und empirische Voraussetzungen einer Teilnahme am moralischen Diskurs. In F. Oser, R. Fatke & O. Höffe (Hrsg.), Transformation und Entwicklung (S. 124-148). Frankfurt a.M.: Suhrkamp.

Kohlberg, L. (1981). Essays on moral development, Bd. 1: The philosophy of moral development. San Francisco: Harper and Row.

Kohlberg, L. (1984): Essays on moral development, Bd. 2: The psychology of moral development. San Francisco: Harper and Row.

Kramer, R. (1968). Moral development in young adulthood. Unpublished doctoral dissertation. University of Chicago.

Krebs, D. L., Vermeulen, S.C.A., Carpendale, J. I. & Denton, K. (1991): Structural and situational influences on moral judgment: The interaction between stage and dilemma. In W. Kurtines & J. Gewirtz (Hrsg.), Handbook of moral behavior and development, Bd. 2 (S. 139-169). Hillsdale: Erlbaum.

Lempert, W. (1988). Soziobiografische Bedingungen der Entwicklung moralischer Urteilsfähigkeit. Kölner Zeitschrift für Soziologie und Sozialpsychologie, 40, 62-92.

Logan, R., Snarey, J. & Schrader, D. (1990). Autonomous versus heteronomous moral judgment types. Journal of Cross-Cultural Psychology, 21, 71-89.



- Nisan, M. & Kohlberg, L. (1982). Universality and variation in moral judgment: A longitudinal and cross-sectional study in Turkey. Child Development, 53, 865-876.
- Nunner-Winkler, G. & Sodian, B. (1988). Children's understanding of moral emotions. Child development, 59, 1323-1338.
- Oser, F. & Althof, W. (1992). Moralische Selbstbestimmung: Modelle der Entwicklung und Erziehung im Wertebereich. Stuttgart: Klett-Cotta.
- Pratt, M. W., Golding, G. & Hunter, W. J. (1984). Does morality have a gender? Sex, sex roles, and moral judgment relationship across the adult life span. Merril-Palmer Quaterly, 30, 321-340.
- Rest, J. R. (1983). Morality. In P. H. Mussen (Hrsg.), Handbook of Child psychology, Vol 3., Cognitive development (S. 556-629). New York: Wiley.
- Roetz, H. (1992). Die chinesische Ethik der Achsenzeit. Frankfurt a.M.: Suhrkamp.
- Rophol, G. (1994). Das Risiko im Prinzip Verantwortung. Ethik und Sozialwissenschaft, 5, 109-120.
- Scott, M. B. & Lyman, S. M. (1976). Praktische Erklärungen. In M. Auwärter, E. Kirsch, S. & M. Schröter (Hrsg.), Seminar: Kommunikation, Interaktion, Identität. Frankfurt a.M.: Suhrkamp.
- Selman, R. L. (1984). Interpersonale Verhandlungen. Eine entwicklungstheoretische Analyse. In W. Edelstein & J. Habermas (Hrsg.), Soziale Interaktion und soziales Verstehen (S. 113-166). Frankfurt a.M.: Suhrkamp.
- Selman, R. L. (1984). Zur Entwicklung interpersonalen Verstehens. Frankfurt a.M.: Suhrkamp.
- Shweder, R. A., Mahapatra M. & Miller, J.G. (1987). Culture and Moral Development. In J. Kagan & S. Lamb (Hrsg.), The emergence of morality in young children., (S. 1 - 83). Chicago: The University of Chicago Press.
- Snarey, J. R. (1985). Cross-cultural universality of social-moral development: A critical review of Kohlbergian research. Psychological Bulletin, 97, 202-232.
- Snarey, J. R. & Keljo, K. (1991). In a Gemeinschaft voice: The cross cultural expansion of moral developmental theory. In W. Kurtines & J. Gewirtz (Hrsg.), Handbook of moral behavior and development, Bd. 1 (S. 395-424). Hillsdale: Earlbaum.
- Sykes, G. M. & Matza, D. (1957). Techniques of neutralization: a theory of delinquency. American Sociological Review, 22, S. 664-670.
- Teo, T., Becker, G. & Edelstein, W. (1995). Variability in Structured Wholeness: Context Factors in L. Kohlberg's Data on the Development of Moral Judgment. Merrill-Palmer Quaterly, 41, 381-391.
- Turiel, E. (1983). The development of social knowledge. Morality and convention. Cambridge, England: Cambridge University Press.

- Vasudev, J. (1986). Kohlbergs Universalitätsanspruch aus indischer Sicht. In: W. Edelstein & G. Nunner-Winkler (Hrsg.), *Zur Bestimmung der Moral* (S. 145-177). Frankfurt a.M.: Suhrkamp.
- Walker, L. J. (1988). The development of moral reasoning. In R. Vasta (Hrsg.), *Annals of Child Development*, Bd. 5, S. 33-78. Greenwich: JAJ Press Inc.
- Walker, L. J. (1989). A longitudinal Study of moral reasoning. *Child Development*, 60, 157-166.
- Walker, L. J. & Taylor, J. H. (1991). Stage transitions in moral reasoning: A longitudinal study of developmental processes. *Developmental Psychology*, 27, 330-337.

## I. Reihe STUDIEN UND BERICHTE

Beim Max-Planck-Institut für Bildungsforschung erhältliche Bände  
(nicht über den Buchhandel beziehbar)

- 62 Andreas Maercker  
**Existenzielle Konfrontation.**  
Eine Untersuchung im Rahmen eines  
psychologischen Weisheitsparadigmas.  
170 S. Erschienen 1995.  
ISBN 3-87985-045-3 DM 19,-
- 61 Alexandra M. Freund  
**Die Selbstdefinition alter Menschen.**  
Inhalt, Struktur und Funktion.  
251 S. Erschienen 1995.  
ISBN 3-87985-057-7 DM 17,-
- 60 Klaus Schömann  
**The Dynamics of Labor Earnings over the Life Course.**  
A Comparative and Longitudinal Analysis of  
Germany and Poland.  
188 S. Erschienen 1994.  
ISBN 3-87985-056-9 DM 13,-
- 59 Frieder R. Lang  
**Die Gestaltung informeller Hilfebeziehungen  
im hohen Alter – Die Rolle von Elternschaft  
und Kinderlosigkeit.**  
Eine empirische Studie zur sozialen Unterstützung  
und deren Effekt auf die erlebte soziale Einbindung.  
177 S. Erschienen 1994.  
ISBN 3-87985-055-0 DM 13,-
- 58 Ralf Th. Krampe  
**Maintaining Excellence.**  
Cognitive-Motor Performance in Pianists  
Differing in Age and Skill Level.  
194 S. Erschienen 1994.  
ISBN 3-87985-054-2 DM 14,-
- 57 Ulrich Mayr  
**Age-Based Performance Limitations in Figural  
Transformations.**  
The Effect of Task Complexity and Practice.  
172 S. Erschienen 1993.  
ISBN 3-87985-053-4 DM 13,-
- 56 Marc Szydlik  
**Arbeitseinkommen und Arbeitsstrukturen.**  
Eine Analyse für die Bundesrepublik Deutschland  
und die Deutsche Demokratische Republik.  
255 S. Erschienen 1993.  
ISBN 3-87985-052-6 DM 21,-
- 55 Bernd Schellhas  
**Die Entwicklung der Ängstlichkeit in Kindheit  
und Jugend.**  
Befunde einer Längsschnittstudie über die  
Bedeutung der Ängstlichkeit für die Entwicklung  
der Kognition und des Schulerfolgs.  
205 S. Erschienen 1993.  
ISBN 3-87985-051-8 DM 13,-
- 54 Falk Fabich  
**Forschungsfeld Schule: Wissenschaftsfreiheit,  
Individualisierung und Persönlichkeitsrechte.**  
Ein Beitrag zur Geschichte  
sozialwissenschaftlicher Forschung.  
235 S. Erschienen 1993.  
ISBN 3-87985-050-X DM 22,-
- 53 Helmut Köhler  
**Bildungsbeteiligung und Sozialstruktur in der  
Bundesrepublik.**  
Zu Stabilität und Wandel der Ungleichheit von  
Bildungschancen.  
133 S. Erschienen 1992.  
ISBN 3-87985-049-6 DM 10,-
- 52 Ulman Lindenberger  
**Aging, Professional Expertise, and Cognitive  
Plasticity.**  
The Sample Case of Imagery-Based Memory  
Functioning in Expert Graphic Designers.  
130 S. Erschienen 1991.  
ISBN 3-608-98257-4 DM 11,-
- 51 Volker Hofmann  
**Die Entwicklung depressiver Reaktionen in  
Kindheit und Jugend.**  
Eine entwicklungspsychopathologische Längs-  
schnittuntersuchung.  
197 S. Erschienen 1991.  
ISBN 3-608-98256-6 DM 14,-
- 50 Georgios Papastefanou (vergriffen)  
**Familiengründung im Lebensverlauf.**  
Eine empirische Analyse sozialstruktureller Bedin-  
gungen der Familiengründung bei den Kohorten  
1929–31, 1939–41 und 1949–51.  
185 S. Erschienen 1990.  
ISBN 3-608-98255-8 DM 15,-

## I. Reihe STUDIEN UND BERICHTE (Fortsetzung)

- 49 Jutta Allmendinger  
**Career Mobility Dynamics.**  
A Comparative Analysis of the United States,  
Norway, and West Germany.  
169 S. Erschienen 1989.  
ISBN 3-608-98254-X DM 13,-
- 48 Doris Sowarka  
**Weisheit im Kontext von Person, Situation und  
Handlung.**  
Eine empirische Untersuchung alltagspsycholo-  
gischer Konzepte alter Menschen.  
275 S. Erschienen 1989.  
ISBN 3-608-98253-1 DM 20,-
- 47 Ursula M. Staudinger  
**The Study of Live Review.**  
An Approach to the Investigation of Intellectual  
Development Across the Life Span.  
211 S. Erschienen 1989.  
ISBN 3-608-98252-3 DM 19,-
- 46 Detlef Oesterreich  
**Die Berufswahlentscheidung von jungen Lehrern.**  
115 S. Erschienen 1987.  
ISBN 3-608-98251-5 DM 9,-
- 45 Hans-Peter Füssel  
**Elternrecht und Schule.**  
Ein Beitrag zum Umfang des Elternrechts in der  
Schule für Lernbehinderte.  
501 S. Erschienen 1987.  
ISBN 3-608-98249-3 DM 22,-
- 44 Diether Hopf  
**Herkunft und Schulbesuch ausländischer Kinder.**  
Eine Untersuchung am Beispiel griechischer Schüler.  
114 S. Erschienen 1987.  
ISBN 3-608-98248-5 DM 8,-
- 43 Eberhard Schröder  
**Entwicklungssequenzen konkreter Operationen.**  
Eine empirische Untersuchung individueller Ent-  
wicklungsverläufe der Kognition.  
112 S. Erschienen 1986.  
ISBN 3-608-98247-7 DM 13,-

Die nicht aufgeführten Bände sind vergriffen,  
bzw. nur noch in Restexemplaren erhältlich.

## II. Reihe MATERIALIEN AUS DER BILDUNGSFORSCHUNG

Beim Max-Planck-Institut für Bildungsforschung erhältliche Bände  
(nicht über den Buchhandel beziehbar)

- 55 Siegfried Reuss und Günter Becker  
**Evaluation des Ansatzes von Lawrence Kohlberg zur Entwicklung und Messung moralischen Urteilens.**  
Immanente Kritik und Weiterentwicklung  
112 S. Erschienen 1996  
ISBN 3-87985-048-8 DM 13,-
- 54 Beate Kraus und Luitgard Trommer  
**Akademiker-Beschäftigung.**  
Sonderauswertung aus der Volkszählung 1987.  
324 S. Erschienen 1995.  
ISBN 3-87985-047-X DM 33,-
- 53 Marianne Müller-Brettel  
**Frieden und Krieg in der psychologischen Forschung.**  
Historische Entwicklungen, Theorien und Ergebnisse.  
296 S. Erschienen 1995.  
ISBN 3-87985-046-1 DM 32,-
- 52 Harald Uhlendorff  
**Soziale Integration in den Freundeskreis.**  
Eltern und ihre Kinder.  
130 S. Erschienen 1995.  
ISBN 3-87985-044-5 DM 15,-
- 51 Peter M. Roeder und Bernhard Schmitz  
**Der vorzeitige Abgang vom Gymnasium.**  
Teilstudie I: Schulformwechsel vom Gymnasium in den Klassen 5 bis 10.  
Teilstudie II: Der Abgang von der Sekundarstufe I.  
159 S. Erschienen 1995.  
ISBN 3-87985-043-7 DM 18,-
- 50 Hannah Brückner  
**Surveys Don't Lie, People Do?**  
An Analysis of Data Quality in a Retrospective Life Course Study.  
86 S. Erschienen 1995.  
ISBN 3-87985-042-9 DM 7,-
- 49 Todd D. Little, Gabriele Oettingen, and Paul B. Baltes  
**The Revised Control, Agency, and Means-ends Interview (CAMI).**  
A Multi-Cultural Validity Assessment Using Mean and Covariance Structures (MACS) Analyses.  
97 S. Erschienen 1995.  
ISBN 3-87985-041-0 DM 8,-
- 48 Hannah Brückner und Karl Ulrich Mayer  
**Lebensverläufe und gesellschaftlicher Wandel.**  
Konzeption, Design und Methodik der Erhebung von Lebensverläufen der Geburtsjahrgänge 1954–1956 und 1959–1961.  
Teil I, Teil II, Teil III.  
169 S., 224 S., 213 S.  
Erschienen 1995.  
ISBN 3-87985-039-9 DM 48,-
- 47 Jochen Fuchs  
**Die bundesdeutschen UNESCO-Projekt-Schulen und ihre internationalen Kontakte und Aktivitäten.**  
57 S. Erschienen 1995.  
ISBN 3-87985-038-0 DM 7,-
- 46 Ursula M. Staudinger, Jacqui Smith und Paul B. Baltes  
**Handbuch zur Erfassung von weisheitsbezogenem Wissen.**  
89 S. Deutsche Ausgabe DM 10,-  
**Manual for the Assessment of Wisdom-Related Knowledge.**  
83 S. Englische Ausgabe  
Erschienen 1994.  
ISBN 3-87985-037-2 DM 10,-
- 45 Jochen Fuchs  
**Internationale Kontakte im schulischen Sektor.**  
Zur Entwicklung und Situation des Schüleraustausches sowie von Schulpartnerschaften in der BRD.  
174 S. Erschienen 1993.  
ISBN 3-87985-035-6 DM 19,-
- 44 Erika Brückner  
**Lebensverläufe und gesellschaftlicher Wandel.**  
Konzeption, Design und Methodik der Erhebung von Lebensverläufen der Geburtsjahrgänge 1919–1921.  
Teil I, Teil II, Teil III, Teil IV, Teil V.  
235 S., 380 S., 200 S., 230 S., 141 S.  
Erschienen 1993.  
ISBN 3-87985-033-X DM 84,-
- 43 Ernst-H. Hoff und Hans-Uwe Hohner  
**Methoden zur Erfassung von Kontrollbewußtsein.**  
Textteil; Anhang.  
99 S. und 178 S. Erschienen 1992.  
ISBN 3-87985-032-1 DM 25,-

## II. Reihe MATERIALIEN AUS DER BILDUNGSFORSCHUNG (Fortsetzung)

- 42 Michael Corsten und Wolfgang Lempert  
**Moralische Dimensionen der Arbeitssphäre.**  
Literaturbericht, Fallstudien und Bedingungs-  
analysen zum betrieblichen und beruflichen  
Handeln und Lernen.  
367 S. Erschienen 1992.  
ISBN 3-87985-031-3 DM 20,-
- 41 Armin Triebel  
**Zwei Klassen und die Vielfalt des Konsums.**  
Haushaltsbudgetierung bei abhängig Erwerbs-  
tätigen in Deutschland im ersten Drittel des  
20. Jahrhunderts.  
Teil I, Teil II.  
416 S., 383 S. Erschienen 1991.  
ISBN 3-87985-030-5 DM 48,-
- 40 Hans-Peter Füssel und Achim Leschinsky (Hrsg.)  
**Reform der Schulverfassung.**  
Wieviel Freiheit braucht die Schule?  
Wieviel Freiheit verträgt die Schule?  
117 S. Erschienen 1991.  
ISBN 3-87985-029-1 DM 13,-
- 39 Gundel Schümer  
**Medieneinsatz im Unterricht.**  
Bericht über Ziel, Anlage und Durchführung einer  
Umfrage in allgemeinbildenden Schulen.  
230 S. Erschienen 1991.  
ISBN 3-87985-025-9 DM 24,-
- 38 Clemens Tesch-Römer  
**Identitätsprojekte und Identitätstransforma-  
tionen im mittleren Erwachsenenalter.**  
312 S. Erschienen 1990.  
ISBN 3-87985-026-7 DM 25,-
- 37 Helmut Köhler  
**Neue Entwicklungen des relativen Schul- und  
Hochschulbesuchs.**  
Eine Analyse der Daten für 1975 bis 1978.  
138 S. Erschienen 1990.  
ISBN 3-87985-024-0 DM 10,-
- 36 Wilfried Spang und Wolfgang Lempert  
**Analyse moralischer Argumentationen.**  
Beschreibung eines Auswertungsverfahrens.  
Textteil: Grundlagen, Prozeduren, Evaluation.  
Anhang: Interviewleitfaden, Tonbandtranskript  
und Auswertungsbeispiele.  
102 und 191 S. Erschienen 1989. DM 29,-
- 35 Karl Ulrich Mayer und Erika Brückner  
**Lebensverläufe und Wohlfahrtsentwicklung.**  
Konzeption, Design und Methodik der Erhebung  
von Lebensverläufen der Geburtsjahrgänge  
1929-1931, 1939-1941, 1949-1951.  
Teil I, Teil II, Teil III.  
261 S., unpaginiert, 175 S.  
Erschienen 1989. DM 39,-
- 34 Christoph Droß und Wolfgang Lempert  
**Untersuchungen zur Sozialisation in der Arbeit  
1977 bis 1988.**  
Ein Literaturbericht.  
204 S. Erschienen 1988. DM 12,-
- 32 Friedrich Edding (Hrsg.)  
**Bildung durch Wissenschaft in neben- und  
nachberuflichen Studien.**  
Tagungsbericht.  
157 S. Erschienen 1988. DM 11,-
- 31 Ellen A. Skinner, Michael Chapman and  
Paul B. Baltes  
**The Control, Agency, and Means-Ends Beliefs  
Interview.**  
A New Measure of Perceived Control in Children  
(School Domain).  
Ein neues Meßinstrument für Kontrollüberzeugun-  
gen bei Kindern (Bereich Schule).  
54 S. Erschienen 1988. DM 9,-
- 29 Ulrich Trommer  
**Aufwendungen für Forschung und Entwicklung  
in der Bundesrepublik Deutschland 1965 bis  
1983.**  
Theoretische und empirisch-statistische Probleme.  
321 S. Erschienen 1987. DM 32,-

Die nicht aufgeführten Bände sind vergriffen,  
bzw. nur noch in Restexemplaren erhältlich.

### III. Einzelpublikationen

#### Beim Max-Planck-Institut für Bildungsforschung erhältliche Titel (nicht über den Buchhandel beziehbar)

Ingo Richter, Peter M. Roeder, Hans-Peter Füssel (Eds.)  
**Pluralism and Education.**  
Current World Trends in Policy, Law, and  
Administration.  
345 S. Berkeley: University of California/USA, 1995.  
DM 25.-

Max-Planck-Institut für Bildungsforschung (Hrsg.)  
**Bekenntnis und Dienst.**  
Reden zum 80. Geburtstag von Dietrich Goldschmidt.  
96 S. Berlin: Max-Planck-Institut für Bildungsforschung, 1995.  
ISBN 3-87985-040-2

Max-Planck-Institut für Bildungsforschung (Hrsg.)  
**Abschied von Hellmut Becker.**  
Reden auf der Trauerfeier am 18. Januar 1994.  
47 S. Berlin: Max-Planck-Institut für Bildungsforschung, 1994.  
ISBN 3-87985-036-4

Max-Planck-Institut für Bildungsforschung (Hrsg.)  
**Bildungsforschung und Bildungspolitik.**  
Reden zum 80. Geburtstag von Hellmut Becker.  
98 S. Berlin: Max-Planck-Institut für Bildungsforschung, 1993.  
ISBN 3-87985-034-8

Wolfgang Schneider and Wolfgang Edelstein (Eds.)  
**Inventory of European Longitudinal Studies in the Behavioral and Medical Sciences.**  
A Project Supported by the European Science Foundation.  
557 S. Munich: Max Planck Institute for Psychological Research, and Berlin: Max Planck Institute for Human Development and Education, 1990.  
ISBN 3-87985-028-3  
DM 58,-

Max-Planck-Institut für Bildungsforschung (Hrsg.)  
**Entwicklung und Lernen.**  
Beiträge zum Symposium anlässlich des 60. Geburtstages von Wolfgang Edelstein.  
98 S. Berlin: Max-Planck-Institut für Bildungsforschung, 1990.  
ISBN 3-87985-023-2

Max-Planck-Institut für Bildungsforschung (Hrsg.)  
**Normative Voraussetzungen und ethische Implikationen sozialwissenschaftlicher Forschung.**  
Beiträge zum Symposium anlässlich des 75. Geburtstages von Dietrich Goldschmidt.  
108 S. Berlin: Max-Planck-Institut für Bildungsforschung, 1990.  
ISBN 3-87985-027-5

Max-Planck-Institut für Bildungsforschung (Hrsg.)  
**25 Jahre Max-Planck-Institut für Bildungsforschung.**  
Festvorträge.  
48 S. Berlin: Max-Planck-Institut für Bildungsforschung, 1989.

Friedrich Edding  
**Mein Leben mit der Politik.**  
126 S. Berlin: Max-Planck-Institut für Bildungsforschung, 1989.

Max-Planck-Institut für Bildungsforschung (Hrsg.)  
**Gewerbliche Unternehmen als Bildungsträger.**  
Beiträge zum Symposium anlässlich des 80. Geburtstages von Friedrich Edding.  
126 S. Berlin: Max-Planck-Institut für Bildungsforschung, 1989.

#### Weitere Schriftenreihen aus dem Max-Planck-Institut für Bildungsforschung (nicht über den Buchhandel erhältlich)

- Beiträge aus dem Forschungsbereich Entwicklung und Sozialisation  
(bitte Liste der Veröffentlichungen anfordern)
- Beiträge aus dem Forschungsbereich Schule und Unterricht  
(bitte Liste der Veröffentlichungen anfordern)
- Literatur-Informationen aus der Bildungsforschung  
(monatliche Neuerwerbungen der Bibliothek; Abonnement DM 60,-/Jahr)

**IV. Buchveröffentlichungen bei Verlagen**  
**(nach dem Erscheinungsjahr geordnet, nur lieferbare Titel;**  
**nur über den Buchhandel zu beziehen)**

Martin Diewald, Karl Ulrich Mayer (Hrsg.)  
**Zwischenbilanz der Wiedervereinigung.**  
 Strukturwandel und Mobilität im  
 Transformationsprozess.  
 352 S. Leverkusen: Leske+Budrich, 1996.

Johannes Huinink, Karl Ulrich Mayer u.a.  
**Kollektiv und Eigensinn.**  
 Lebensverläufe in der DDR und danach.  
 414 S. Berlin: Akademie Verlag, 1995.

Johannes Huinink  
**Warum noch Familie?**  
 Zur Attraktivität von Partnerschaft und Elternschaft in  
 unserer Gesellschaft.  
 385 S. Frankfurt a. M./New York: Campus, 1995.

Heike Trappe  
**Emanzipation oder Zwang?**  
 Frauen in der DDR zwischen Beruf, Familie und  
 Sozialpolitik.  
 242 S. Berlin: Akademie Verlag, 1995.

Heike Solga  
**Auf dem Weg in eine klassenlose Gesellschaft?**  
 Klassenlagen und Mobilität zwischen Generationen in  
 der DDR.  
 265 S. Berlin: Akademie Verlag, 1995.

Lothar Krappmann und Hans Oswald  
**Alltag der Schulkinder.**  
 Beobachtungen und Analysen von Interaktionen und  
 Sozialbeziehungen.  
 224 S. Weinheim/München: Juventa, 1995.

Freya Dittmann-Kohli  
**Das persönliche Sinnsystem.**  
 Ein Vergleich zwischen frühem und spätem  
 Erwachsenenalter.  
 402 S. Göttingen/Bern/Toronto/Seattle: Hogrefe, 1995.

Hartmut Zeiher und Helga Zeiher  
**Orte und Zeiten der Kinder.**  
 Soziales Leben im Alltag von Großstadtkindern.  
 223 S. Weinheim/München: Juventa, 1994.

Christiane Lange-Küttner  
**Gestalt und Konstruktion.**  
 Die Entwicklung der grafischen Kompetenz beim  
 Kind.  
 242 S. Bern/Toronto: Huber, 1994.

Jutta Allmendinger  
**Lebensverlauf und Sozialpolitik.**  
 Die Ungleichheit von Mann und Frau und ihr  
 öffentlicher Ertrag.  
 302 S. Frankfurt a. M./New York: Campus, 1994.

Wolfgang Lauterbach  
**Berufsverläufe von Frauen.**  
 Erwerbstätigkeit, Unterbrechung und Wiedereintritt.  
 289 S. Frankfurt a. M./New York: Campus, 1994.

Arbeitsgruppe Bildungsbericht am  
 Max-Planck-Institut für Bildungsforschung  
**Das Bildungswesen in der Bundesrepublik  
 Deutschland.**  
 Strukturen und Entwicklungen im Überblick.  
 843 S. Reinbek: Rowohlt, 1994 (4., vollständig über-  
 arbeitete und erweiterte Neuauflage).

Hellmut Becker und Gerhard Kluchert  
**Die Bildung der Nation.**  
 Schule, Gesellschaft und Politik vom Kaiserreich zur  
 Weimarer Republik.  
 538 S. Stuttgart: Klett-Cotta, 1993.

Rolf Becker  
**Staatsexpansion und Karrierechancen.**  
 Berufsverläufe im öffentlichen Dienst und in der  
 Privatwirtschaft.  
 303 S. Frankfurt a.M./New York: Campus, 1993.

Wolfgang Edelstein und  
 Siegfried Hoppe-Graff (Hrsg.)  
**Die Konstruktion kognitiver Strukturen.  
 Perspektiven einer konstruktivistischen  
 Entwicklungspsychologie.**  
 328 S. Bern/Stuttgart/Toronto: Huber, 1993.

Wolfgang Edelstein, Gertrud Nunner-Winkler  
 und Gil Noam (Hrsg.)  
**Moral und Person.**  
 418 S. Frankfurt a.M.: Suhrkamp, 1993.

Lothar Lappe  
**Berufsperspektiven junger Facharbeiter.**  
 Eine qualitative Längsschnittanalyse zum Kernbereich  
 westdeutscher Industriearbeit.  
 394 S. Frankfurt a.M./New York: Campus, 1993.

Detlef Oesterreich  
**Autoritäre Persönlichkeit und Gesellschaftsordnung.**  
 Der Stellenwert psychischer Faktoren für politische  
 Einstellungen – eine empirische Untersuchung von  
 Jugendlichen in Ost und West.  
 243 S. Weinheim/München: Juventa, 1993.



## IV. Buchveröffentlichungen bei Verlagen (Fortsetzung)

Marianne Müller-Brettel

### **Bibliographie Friedensforschung und Friedenspolitik:**

Der Beitrag der Psychologie 1900–1991.  
(Deutsch/Englisch)

383 S. München/London/New York/Paris: Saur, 1993.  
Paul B. Baltes und Jürgen Mittelstraß (Hrsg.)

### **Zukunft des Alterns und gesellschaftliche Entwicklung.**

(= Forschungsberichte der Akademie der Wissenschaften zu Berlin, 5.)

814 S. Berlin/New York: De Gruyter, 1992.

Matthias Grundmann

### **Familienstruktur und Lebensverlauf.**

Historische und gesellschaftliche Bedingungen individueller Entwicklung.

226 S. Frankfurt a.M./New York: Campus, 1992.

Karl Ulrich Mayer (Hrsg.)

### **Generationsdynamik in der Forschung.**

245 S. Frankfurt a.M./New York: Campus, 1992.

Erika M. Hoerning

### **Zwischen den Fronten.**

Berliner Grenzgänger und Grenzhändler 1948–1961.

266 S. Köln/Weimar/Wien: Böhlau, 1992.

Ernst-H. Hoff

### **Arbeit, Freizeit und Persönlichkeit.**

Wissenschaftliche und alltägliche Vorstellungsmuster.

238 S. Heidelberg: Asanger Verlag, 1992 (2. überarbeitete und aktualisierte Auflage).

Erika M. Hoerning

### **Biographieforschung und Erwachsenenbildung.**

223 S. Bad Heilbrunn: Klinkhardt, 1991.

Max-Planck-Institut für Bildungsforschung

### **Traditions et transformations.**

Le système d'éducation en République fédérale d'Allemagne.

341 S. Paris: Economica, 1991.

Dietrich Goldschmidt

### **Die gesellschaftliche Herausforderung der Universität.**

Historische Analysen, internationale Vergleiche, globale Perspektiven.

297 S. Weinheim: Deutscher Studien Verlag, 1991.

Uwe Henning und Achim Leschinsky (Hrsg.)

### **Enttäuschung und Widerspruch.**

Die konservative Position Eduard Sprangers im Nationalsozialismus. Analysen – Texte – Dokumente.

213 S. Weinheim: Deutscher Studien Verlag, 1991.

Ernst-H. Hoff, Wolfgang Lempert und Lothar Lappe  
**Persönlichkeitsentwicklung in Facharbeiterbiographien.**

282 S. Bern/Stuttgart/Toronto: Huber, 1991.

Karl Ulrich Mayer, Jutta Allmendinger und Johannes Huinink (Hrsg.)

### **Vom Regen in die Traufe: Frauen zwischen Beruf und Familie.**

483 S. Frankfurt a.M./New York: Campus, 1991.

Maria von Salisch

### **Kinderfreundschaften.**

Emotionale Kommunikation im Konflikt.

153 S. Göttingen/Toronto/Zürich: Hogrefe, 1991.

Paul B. Baltes and Margret M. Baltes (Eds.)

### **Successful Aging: Perspectives from the Behavioral Sciences.**

397 pp. Cambridge: Cambridge University Press, 1990.

Paul B. Baltes, David L. Featherman and Richard M. Lerner (Eds.)

### **Life-Span Development and Behavior.**

368 pp. Vol. 10. Hillsdale, N.J.: Erlbaum, 1990.

Achim Leschinsky and Karl Ulrich Mayer (Eds.)

### **The Comprehensive School Experiment Revisited: Evidence from Western Europe.**

211 pp. Frankfurt a.M./Bern/New York/Paris: Lang 1990.

Karl Ulrich Mayer (Hrsg.)

### **Lebensverläufe und sozialer Wandel.**

467 S. Opladen: Westdeutscher Verlag, 1990.

(= Sonderheft 31 der KZfSS).

Karl Ulrich Mayer and Nancy Brandon Tuma (Eds.)

### **Event History Analysis in Life Course Research.**

320 pp. Madison, Wis.: The University of Wisconsin Press, 1990.

Hans J. Nissen, Peter Damerow und Robert K. Englund  
**Frühe Schrift und Techniken der Wirtschaftsverwaltung im alten Vorderen Orient.**

Informationsspeicherung und -verarbeitung vor 5000 Jahren.

Katalog zur gleichnamigen Ausstellung Berlin-Charlottenburg, Mai–Juli 1990.

222 S. Bad Salzdetfurth: Franzbecker, 1990.  
(2. Aufl. 1991).

Peter Alheit und Erika M. Hoerning (Hrsg.)

### **Biographisches Wissen.**

Beiträge zu einer Theorie lebensgeschichtlicher Erfahrung.

284 S. Frankfurt a.M./New York: Campus, 1989.

#### IV. Buchveröffentlichungen bei Verlagen (Fortsetzung)

Arbeitsgruppe am Max-Planck-Institut für  
Bildungsforschung

**Das Bildungswesen in der Bundesrepublik  
Deutschland.**

Ein Überblick für Eltern, Lehrer und Schüler.  
Japanische Ausgabe: 348 S. Tokyo: Toshindo  
Publishing Co. Ltd., 1989.

Hans-Peter Blossfeld

**Kohortendifferenzierung und Karriereprozeß.**

Eine Längsschnittstudie über die Veränderung der  
Bildungs- und Berufschancen im Lebenslauf.  
185 S. Frankfurt a.M./New York: Campus, 1989.

Hans-Peter Blossfeld, Alfred Hamerle and  
Karl Ulrich Mayer

**Event History Analysis.**

Statistical Theory and Application in the Social  
Sciences.  
297 pp. Hillsdale, N.J.: Erlbaum, 1989.

Erika M. Hoerning und Hans Tietgens (Hrsg.)  
**Erwachsenenbildung: Interaktion mit der  
Wirklichkeit.**

200 S. Bad Heilbrunn: Klinkhardt, 1989.

Johannes Huinink

**Mehrebenensystem-Modelle in den Sozialwissen-  
schaften.**

292 S. Wiesbaden: Deutscher Universitäts-Verlag,  
1989.

Kurt Kreppner and Richard M. Lerner (Eds.)

**Family Systems and Life-Span Development.**

416 pp. Hillsdale, N.J.: Erlbaum, 1989.

Bernhard Schmitz

**Einführung in die Zeitreihenanalyse.**

Modelle, Softwarebeschreibung, Anwendungen.  
235 S. Bern/Stuttgart/Toronto: Huber, 1989.

Eberhard Schröder

**Vom konkreten zum formalen Denken.**

Individuelle Entwicklungsverläufe von der Kindheit  
zum Jugendalter.

328 S. Bern/Stuttgart/Toronto: Huber, 1989.

Michael Wagner

**Räumliche Mobilität im Lebensverlauf.**

Eine empirische Untersuchung sozialer Bedingungen  
der Migration.

226 S. Stuttgart: Enke, 1989.

Paul B. Baltes, David L. Featherman and  
Richard M. Lerner (Eds.)

**Life-Span Development and Behavior.**

338 pp. Vol. 9. Hillsdale, N.J.: Erlbaum, 1988.

Paul B. Baltes, David L. Featherman and  
Richard M. Lerner (Eds.)

**Life-Span Development and Behavior.**

337 pp. Vol. 8. Hillsdale, N.J.: Erlbaum, 1988.

Lothar Krappmann

**Soziologische Dimensionen der Identität.**

Strukturelle Bedingungen für die Teilnahme an  
Interaktionsprozessen.

231 S. Stuttgart: Klett-Cotta, 7. Aufl., 1988

(= Standardwerke der Psychologie).

Detlef Oesterreich

**Lehrerkooperation und Lehrersozialisation.**

159 S. Weinheim: Deutscher Studien Verlag, 1988.

Michael Bochow und Hans Joas

**Wissenschaft und Karriere.**

Der berufliche Verbleib des akademischen Mittelbaus.  
172 und 37 S. Frankfurt a.M./New York: Campus, 1987.

Hans-Uwe Hohner

**Kontrollbewußtsein und berufliches Handeln.**

Motivationale und identitätsbezogene Funktionen  
subjektiver Kontrollkonzepte.

201 S. Bern/Stuttgart/Toronto: Huber, 1987.

Bernhard Schmitz

**Zeitreihenanalyse in der Psychologie.**

Verfahren zur Veränderungsmessung und Prozeß-  
diagnostik.

304 S. Weinheim/Basel: Deutscher Studien Verlag/  
Beltz, 1987.

Margret M. Baltes and Paul B. Baltes (Eds.)

**The Psychology of Control and Aging.**

415 pp. Hillsdale, N.J.: Erlbaum, 1986.

Paul B. Baltes, David L. Featherman and  
Richard M. Lerner (Eds.)

**Life-Span Development and Behavior.**

334 pp. Vol. 7. Hillsdale, N.J.: Erlbaum, 1986.

Hans-Peter Blossfeld, Alfred Hamerle and  
Karl Ulrich Mayer

**Ereignisanalyse.**

Statistische Theorie und Anwendung in den  
Wirtschafts- und Sozialwissenschaften.

290 S. Frankfurt a.M./New York: Campus, 1986.

## IV. Buchveröffentlichungen bei Verlagen (Fortsetzung)

Axel Funke, Dirk Hartung, Beate Kraus und Reinhard Nuthmann

**Karrieren außer der Reihe.**

Bildungswege und Berufserfolge von Stipendiaten der gewerkschaftlichen Studienförderung.  
256 S. Köln: Bund, 1986.

Ernst-H. Hoff, Lothar Lappe und Wolfgang Lempert (Hrsg.)

**Arbeitsbiographie und Persönlichkeitsentwicklung.**

288 S. Bern/Stuttgart/Toronto: Huber, 1986.

Klaus Hüfner, Jens Naumann, Helmut Köhler und Gottfried Pfeffer

**Hochkonjunktur und Flaute: Bildungspolitik in der Bundesrepublik Deutschland 1967–1980.**

361 S. Stuttgart: Klett-Cotta, 1986.

Jürgen Staube

**Parlamentsvorbehalt und Delegationsbefugnis.**

Zur „Wesentlichkeitstheorie“ und zur Reichweite legislativer Regelungskompetenz, insbesondere im Schulrecht.

419 S. Berlin: Duncker & Humblot, 1986.

Hans-Peter Blossfeld

**Bildungsexpansion und Berufschancen.**

Empirische Analysen zur Lage der Berufsanfänger in der Bundesrepublik.

191 S. Frankfurt a.M./New York: Campus, 1985.

Christel Hopf, Knut Nevermann und Ingrid Schmidt

**Wie kamen die Nationalsozialisten an die Macht.**

Eine empirische Analyse von Deutungen im Unterricht.

344 S. Frankfurt a.M./New York: Campus, 1985.

John R. Nesselroade and Alexander von Eye (Eds.)

**Individual Development and Social Change:**

**Explanatory Analysis.**

380 pp. New York: Academic Press, 1985.

Michael Jenne

**Music, Communication, Ideology.**

185 pp. Princeton, N.J.: Birch Tree Group Ltd., 1984.

Gero Lenhardt

**Schule und bürokratische Rationalität.**

282 S. Frankfurt a.M.: Suhrkamp, 1984.

Achim Leschinsky und Peter Martin Roeder

**Schule im historischen Prozeß.**

Zum Wechselverhältnis von institutioneller Erziehung und gesellschaftlicher Entwicklung.

545 S. Frankfurt a.M./Berlin/Wien: Ullstein, 1983.

Max Planck Institute for Human Development and Education

**Between Elite and Mass Education.**

Education in the Federal Republic of Germany.

348 pp. Albany: State University of New York Press, 1983.

Margit Osterloh

**Handlungsspielräume und Informationsverarbeitung.**

369 S. Bern/Stuttgart/Toronto: Huber, 1983.

Knut Nevermann

**Der Schulleiter.**

Juristische und historische Aspekte zum Verhältnis von Bürokratie und Pädagogik.

314 S. Stuttgart: Klett-Cotta, 1982.

Gerd Sattler

**Englischunterricht im FEGA-Modell.**

Eine empirische Untersuchung über inhaltliche und methodische Differenzierung an Gesamtschulen.

355 S. Stuttgart: Klett-Cotta, 1981.

Christel Hopf, Knut Nevermann und Ingo Richter

**Schulaufsicht und Schule.**

Eine empirische Analyse der administrativen Bedingungen schulischer Erziehung.

428 S. Stuttgart: Klett-Cotta, 1980.

Diether Hopf

**Mathematikunterricht.**

Eine empirische Untersuchung zur Didaktik und Unterrichtsmethode in der 7. Klasse des Gymnasiums.

251 S. Stuttgart: Klett-Cotta, 1980.

Max-Planck-Institut für Bildungsforschung  
Projektgruppe Bildungsbericht (Hrsg.)

**Bildung in der Bundesrepublik Deutschland.**

Daten und Analysen.

Bd. 1: Entwicklungen seit 1950.

Bd. 2: Gegenwärtige Probleme.

1404 S. Stuttgart: Klett-Cotta, 1980.

Dietrich Goldschmidt und Peter Martin Roeder (Hrsg.)

**Alternative Schulen?**

Gestalt und Funktion nichtstaatlicher Schulen im Rahmen öffentlicher Bildungssysteme.

623 S. Stuttgart: Klett-Cotta, 1979.