Moral Development and the Social Environment

Studies in the Philosophy and Psychology of Moral Judgment and Education

Re-publication of selected chapters.

Original page numbers in [red]

1985

Unchanged, selected edition 2000
Acknowledgement

I like to thank Prof. Michael Gross, Precedent Publishing Inc., for holding this book in print for so many years and for his permission for reprint.

No part of this book may be reproduced in any way without the written permission of the copyright holder.

Contact:
E-mail: Georg.Lind@uni-konstanz.de

Further publications by Georg Lind:
www.uni-konstanz.de/ag-moral/b-liste.htm

These articles were originally published in:
Preface to the Unchanged Selected Edition

Georg Lind

Since the early seventies, several research projects on moral development in German speaking countries (Germany, Switzerland and Austria) have been stimulated by the works of Lawrence Kohlberg, Jean Piaget and Emil Durkheim. For many years this research, however, was widely unknown outside these countries because of language barriers. So I was happy that in the early eighties the series editor of Precendent Publishing Inc., Professor Thomas E. Wren, and the publisher, Mr. Jacob Cohen, have invited me to edit the book Moral Development and the Social Environment. Studies in the Psychology and Philosophy of Moral Judgment and Education (originally published in 1985), which should contain articles in English by German speaking scholars on their empirical and philosophical studies.

We – my co-editors, Hans A. Hartmann, professor of psychology at the University of Augsburg, and the late Roland Wakenhut, professor of social and industrial psychology at the Catholic University of Eichstätt and I – asked a wide array of German speaking colleagues, who did excellent research in the field of moral development, to write chapters for this book: Tino Bargel, Hans Bertram, Gerhard Bohm, Rainer Döbert, Horst Heidbrink, Gertrud Nunner-Winkler, Fritz Oser, Gerhard Portele, Johann-Ulrich Sandberger, Michael Schenk, Andre Schläfli, and Rainer Senger. We also wrote several chapters. We were especially happy that the philosopher Jürgen Habermas, permitted to use a chapter from a German book as an introduction into the philosophical debate on Kohlberg’s theory, and that the late Lawrence Kohlberg himself wrote the Foreword to our book.

The book was well received. Unfortunately, it is out of print for many years already, and a second edition does not seem feasible. I am receiving many requests especially for the chapters on the theory and research using the Moral Judgment Test (MJT), which we developed at the University of Konstanz. To make this information accessible again, I decided to re-publish the chapters which I co-authored on the internet, plus the foreword by Lawrence Kohlberg. For this re-publication only minor changes (typographic errors) were made. For her help with corrections and editing, I wish to thank Patricia Knoop.

Since 1985, much new research has been done building this work und using the MJT. See G. Lind Can Ethics be Taught?. Berlin: Logos-Verlag (in German, ISDN 3-89722-255-8) and www.uni-konstanz.de/ag-moral/ for further information, links and publications.

Georg Lind
Konstanz, Nov. 2000

* In meantime renamed as Moral Competence Test (MCT)
Precedent Studies in Ethics and the Moral Sciences

The term “moral sciences,” which is the rubric for the present series, originated with John Stuart Mill, who used it to cover what are now called the social and behavioral sciences. Accordingly, the Precedent Studies in Ethics and the Moral Sciences have as a general subject matter the rich and problematic interspace between philosophical ethics and such empirical but person- and society-oriented disciplines as developmental psychology, sociology, and anthropology, to cite a few of the more salient moral sciences.”

The present volume is about moral judgment, especially its exercise in selected social settings such as the university. Because of this focus, as well as its simultaneous deployment of empirical techniques and philosophical arguments, it vividly instantiates the term “moral sciences” as well as the term Geisteswissenschaften, which was coined in Germany as a counterpart to Mill’s term.

The contributors to this volume are all German or German-speaking psychologists, sociologists, and philosophers of morality, most of whom have collaborated on long-ranged research projects in Europe involving university socialization. Taken singly as well as in combination, their essays make it clear that the study of morality is an empirical as well as a conceptual task, one that invokes data collection, statistical analysis, and the formulation and testing of hypotheses. It may be unreasonable to expect moral philosophers to assume the responsibility for such empirical inquiries in addition to their own theoretical pursuits, but it is not unreasonable to expect them to take a keen interest in the outcomes. For instance, the very question of whether moral cognition admits of measurement is a philosophical as well as scientific issue, the answer to which is to a large extent dependent on the de facto success or failure of the methodologies described in the following pages. Whether moral reasoning has a content-specific domain, whether its structures transcend specific issues of justice, obedience, etc. – these and similar questions suggest that moral philosophers and “moral scientists” have much to say to each other. This volume not only exemplifies that conversation but also makes an important contribution to it.

Thomas E. Wren
General Editor
Contents

Preface to the Unchanged Selected Edition by Georg Lind .............. 3
Precedent Studies in Ethics and Moral Sciences ........................... 4
Preface to the English Edition ............................................. 6
Foreword by Lawrence Kohlberg ........................................... 11

Theory and Methods

1 The Theory of Moral-Cognitive Development:
   A Socio-Psychological Assessment
   by Georg Lind ................................................................. 14
   (pp. 21 - 53)

2 Moral Competence and Democratic Personality
   by Georg Lind, Johann-Ulrich Sandberger and Tino Bargel ....... 45
   (pp. 55 - 77)

3 Testing for Moral Judgment Competence
   by Georg Lind and Roland Wakenhut .................................. 67
   (pp. 79 - 105)

Moral Judgment and Its Social Milieu

4 Attitude Change or Cognitive-Moral Development?
   by Georg Lind ................................................................. 91
   (pp. 173 - 192)

5 Moral Judgment Competence and Political learning
   by Horst Heidbrink ......................................................... 111
   (259 - 271)

References ................................................................. 123

Appendix: Authors of the original edition ................................. 150
Appendix: Contents of the original edition ................................ 155

1 The original page numbers of the book are printed in parentheses (pp. xx – xx).
Contents

Preface ix
Foreword by Lawrence Kohlberg xv

Theory and Methods

Philosophical Notes on Moral Judgment Theory by Jürgen Habermas 3
The Theory of Moral-Cognitive Development: A Socio-Psychological Assessment by Georg Lind 21
Moral Competence and Democratic Personality by Georg Lind, Johann-Ulrich Sandberger, and Tino Bargel 55
Testing for Moral Judgment Competence by Georg Lind and Roland Wakenhut 79

Moral Judgment and Its Social Milieu

Social Inequality and Moral Development by Hans Bertram 109
Value Change and Morality by Rainer Döbert and Gertrud Nunner-Winkler 125
The Thin Line Phenomenon: Helping Bank Trainees Form a Social and Moral Identity by Fritz Oser and André Schläfli 155
Attitude Change or Cognitive-Moral Development? by Georg Lind 173
How Scientists Think about Science and Morality by Gerhard Portele 193
Morality and the Military Life-World by Thomas Krämer-Badoni and Roland Wakenhut 205
Segmentation of Soldiers’ Moral Judgment by Rainer Senger 221
Single-Issue Movements: Political Commitment and Moral Judgment by Michael Schenk and Gerhard Bohm 243
Philosophical Epilogue: Phylogeny and Ontogeny

What Is Social About Morality? Morals, Morality, and Ethics in Social Science Perspectives by Hans A. Hartmann

Contributors 296
References 299
Index of Names 319
Index of Subjects 325
Preface to the English Edition

The contributors to this volume represent diverse disciplines, but they have as their common concern the topic announced in the book's title: the relation and the interaction of individual or group-specific moral development on the one hand, and the social milieu on the other. Although deeply involved in empirical research, the authors maintain that research on moral development can be pursued properly only in conjunction with a well-formulated theory of the relationship between society, cognition, and behavior.

Readers not familiar with recent developments in the German-language literature on social philosophy and psychology, both its theoretical side and its application to educational contexts, will find this book a useful introduction to the increasingly important scene of German social theory as well as a stimulating discussion of the issues themselves. Its original impetus was a series of extensive and still on-going research projects in West Germany and Switzerland, concerning the process of moral development in social institutions such as schools, universities, and the military. In their various essays the authors all return to the same theme: the ability to formulate genuine and coherent moral judgments reflects social conditions at two levels, that of individual socialization and that of the historical development of the social system.

The book's fusion of developmental psychology, sociology, and social psychology is a distinctly German phenomenon, which in the present volume is related directly to the work of Jean Piaget of Switzerland and, more recently, Lawrence Kohlberg of the United States, who has written the foreword to the much-revised English edition of the book. The authors are not uncritical of these two figures, and are especially concerned to correct the imbalances which they perceive in Kohlberg's work between cognition and affect, as well as between the structures and contents of moral judgment. They are all deeply influenced by the cognitive-developmental tradition.

The first of the three major sections of the book deals explicitly with theoretical and methodological questions. Jürgen Habermas, who is one of Kohlberg's most important dialogue partners in Europe, comments on the theory of cognitive moral development from the viewpoint of his own theory of communicative action. His essay (ch. 1), which is part of a larger work on moral psychology, focuses on four questions that are central for the present discussion of Kohlberg's theory: (1) the question about the status of the postconventional stages, (2) doubts about the adequacy of a normative reference point in the face of the empirical regression phenomenon, (3) the problem of fitting relativists and value skeptics into the stage model, and (4) the necessity of including psychodynamic aspects of judgment formation within the structural theory of moral development. Georg Lind (ch. 2) attempts to assess Kohlberg's theory from
a socio-psychological perspective. Regarding the history of psychology, he says that cognitive developmental theory meant a shift from the “external effect” to the “internal structure” model of psychological functioning. From a review of the literature and from his own research he concludes that Kohlberg’s theory of moral cognitive development has enriched social psychology with new and psychologically meaningful data and, at the same time, with empirical hypotheses that are *gehaltvoll* (informative) and empirically valid. However, Lind proposes three broad revisions in the theory: (1) it should be made clear that affect and cognition are analytically clearly distinct but ontologically inseparable aspects of moral judgment behavior; (2) accordingly, moral development should be viewed as a “two-sided” process of change in affects and in the structure of judgment, considered in a more formal sense than Kohlberg has suggested; and (3) the assumption of “structural wholeness” should be regarded not as an empirical hypothesis but rather as a methodological criterion that serves to measure the cognitive aspect of moral development. Finally, he discusses the relation of individual moral judgment and the social environment, including the issue of personalism versus situationism, the impact of what Max Weber called . positional fragmentation” on the mode of moral judgment, and the role of social selection in regard to moral development. Although they are trained in quite different fields, Georg Lind, Johann-Ulrich Sandberger, and Tino Bargel have collaborated on a cross-national longitudinal research about university socialization. In their joint essay (ch. 3) they investigate the notion of a democratic personality from three theoretical points of view, the cognitive-developmental (Piaget, Kohlberg), the attitude-structural (Converse, A. Campbell), and the psychodynamic (Freud, Adorno) viewpoints. In the theoretical part of their paper, Lind, Sandberger, and Bargel argue that between the models of these different approaches there are some intriguing similarities which suggest a common conceptual basis for research. In the empirical part of their essay, they present evidence in support of a rather close relationship between moral judgment competence, ego-strength, and democratic orientations. In studies of the social ecology of moral behavior and thought, the research designs usually become quite complex and thus require rather large samples and easy-to-use questionnaires. Georg Lind and Roland Wakenhut (ch. 4) present a new method of assessing affective and cognitive aspects of moral judgment behavior, that of what they call “Experimental Questionnaires.” These tests combine the transparency and economy of classical attitude questionnaires with the methodological postulates of cognitive developmental theory. To illustrate the methodology which they are proposing, they describe the construction of the Moralisches Urteil Test (MUT) by Lind and the Moralisches Urteil Fragebogen (MUF) by Wakenhut, both of which have been widely used in moral judgment research in Europe. Both instruments have proved especially useful for research on moral development in differen-
tial social milieus. The second and longest section of the book focuses on the social context of moral judgment, with studies of specific social milieus such as the family, the professions, the university, the military, the scientific world, and political subcultures. The authors' major interest was in studying particular facets and levels of the relation between the individual's moral behavior on the one hand and roles, norms, and institutions on the other. Rather than merely trying to quantify the impact of the social environment on the development of moral competence, the essays deal with the quality of this relationship. In the first essay in this section, Hans Bertram (ch. 5) suggests a multi-level model of social reproduction. He presents empirical findings which show a significant impact of social class and the structure of work-organization (level one), mediated by parents' values, personalities, and forms of interaction between the parents (level two), on the moral cognitive development of children (level three). Bertram concludes that in spite of the increasingly greater importance of formalized processes of education, the family still plays an outstanding role in the individual's cognitive-moral socialization. Rainer Döbert and Gertrud Nunner-Winkler (ch. 6), propose the thesis that, contrary to Ronald Inglehart's now widely-accepted theory, value change cannot be reduced to an essentially non-cognitive substitution of one value content for another but, following Max Weber as well as Kohlberg, instead should be regarded as a cognitive process in which traditions are dissolved. They point out that there are three ways of describing the validity of norms, corresponding to (1) their substantive validity, which refers to the existence or nonexistence of norms and their relative place in a value hierarchy; (2) their social validity, which refers to the degree to which small or large groups or the whole society hold such norms; and finally (3) their validity mode, which refers to the legitimacy and justifiability of social norms from the individual's point of view. Döbert and Nunner-Winkler argue that the third of these dimensions of validity is, in spite of frequent neglect, as least as important as the others. They illustrate and support this claim through the analysis of public debates and the interview-based study of male and female adolescents on the issues of abortion and conscientious objection to military service. Fritz Oser and André Schläfli (ch. 7) inquire about conditions and ways of fostering cognitive-moral development during professional training, which in Switzerland as well as in Germany is provided for adolescents in a secondary school curriculum that is an alternative to the college-bound Gymnasium. They show that, contrary to common beliefs, the trainees welcome the opportunity to discuss moral dilemmas and experience an increase of maturity in regard to their values, socio-political attitudes, and moral judgment behavior. Using interview and questionnaire data, Oser and Schläfli argue that moral education is enhanced by the tension and conflict produced by considering real life dilemmas and by walking the "thin line" between milieu and the individual's needs, duties, and rights.
Georg Lind (ch. 8) tries to answer the question of why so many studies concerning the impact of schools and colleges on the individual's socioaffective development produce inconsistent if not negative findings. He hypothesizes that this finding does not reflect the actual effects of education but rather a methodological bias. Because classical research methods (e.g., attitude tests) rest on a simplistic model of man, they fail to assess the cognitive aspects of social, political, or moral attitudes. By reanalyzing findings of classical attitude research as well as data from his own moral judgment studies, Lind shows that higher education consistently stimulates the integration and differentiation of attitudes. Gerhard Portele (ch. 9) analyzes the views of natural and social scientists concerning the aims and means of scientific research and concerning morality. He comes to a conclusion which confirms the fears of those who regard the ever-growing industrialization of our modern times as presaged by the famous Chaplin film of that title. Portele shows that natural scientists and technologists are inclined, much more than are their humanistic and social science counterparts, toward a segmentation between the world of research and everyday life, as well as toward political alienation and conventional or even preconventional moral perspectives in scientific and non-scientific issues. The differential moral thinking of individuals in contrasting social environments is also the topic of the study by Thomas Krämer-Badoni and Roland Wakenhut (ch. 10) of soldiers and reservists of the German Armed Forces, conscientious objectors, and others. They have investigated the effect of the objective and subjective conditions of military life on the moral judgment pattern of its members, concluding that the German Armed Forces are still morally cut off or “segmented” from the civilian part of society. This phenomenon of moral segmentation is studied more extensively by Rainer Senger (ch. 11), who discusses the educational ideal of the “citizen in uniform,” which aims at the integration of the military and the democratic society. On the grounds of a critical discussion of Charles Levine's theory of inconsistency of moral judgment, Senger develops the outlines of his own theory of segmentation, which he illustrates with his interviews with officers and soldiers. Michael Schenk and Gerhard Bohm (ch. 12) employ the cognitive developmental theory of moral judgment to study two political groups, each of which were involved in single-issue movements but in pursuit of quite different aims. Although the two groups differed less in regard to their mean judgment competence than the authors had expected, Schenk and Bohm's data show characteristic differences regarding the distribution of the moral reasoning within each group. While all moral levels were equally present in the group with “egocentric” political aims, in the group with the altruist goals the postconventional reasoning dominated and preconventional thought was more frequent than conventional ones, a difference which resembles Norma Haan's well-known findings concerning ego development
and moral reasoning. Horst Heidbrink (ch. 13) argues that the moral domain and the political domain cannot be reduced on one another and that it is necessary to keep them distinct. Political judgments usually involve processing concrete information and hence require complex cognitive operations, whereas moral judgments can be fruitfully studied even with fictitious and abstract situations. Nevertheless, from his own experimental work Heidbrink concludes that moral judgment competence facilitates political learning. He favors a curriculum for social and political studies which makes use of Kohlberg's moral pedagogy but resists the temptation to consider itself as apolitical. In the epilogue to the book, Hans A. Hartmann (ch. 14) describes the history of morality as a socio-cultural phenomenon and as interwoven with the development of moral philosophy and socio-psychological research on morality. Hartmann puts the research presented in the other essays into historical and philosophical perspective by sketch philosophy: the formalist ethics of Kant, French positivism (Comte, Durkheim), genetic epistemology (Piaget), and the cognitive development theory by Kohlberg as well as the theory of sociocultural evolution developed by Habermas and his colleagues.

Most of the essays originally appeared in this book's German counterpart, *Moralisches Urteilen und soziale Umwelt* (1983), but they have been expanded and substantially revised for their appearance in English as part of the Precedent Studies in Ethics and the Moral Sciences. The editors wish to thank all those who have helped in various ways to make this publication possible. Above all, we thank the authors for their great patience and willingness to revise their original manuscripts and to provide supplementary information for the English-speaking reader. We also wish to thank the Universities of Augsburg and Konstanz as well as the Hochschulsozialisation research project, which in various ways have supported the preparation of this book, as well as the following individuals. For their assistance, in several cases very extensive, in translating the German essays or preparing them for translation, we wish to thank Christian Lenhardt, Denise Deissenberg, Naomi Knapp, James Fearn, Raoul Eshelman, and Professor Bernard Mausner. We wish to thank Doris Lang and Karin Gauggel for keying the manuscripts into the word processor, Leonore Link, Mary D. Hawley, and Sheila Nolan Whalen for their editorial assistance, and Bruce Uttley, Director of the Department of Computer Services at the University of Waterloo, for his patient help in converting our computer files into a format suitable for typesetting. Last but not least we wish to express our gratitude to Thomas Wren, the general editor and translator, for his encouragement and his extensive advice, the outcome of which is not only a translation from the original German but also a substantially revised edition, updated and rendered more suitable to an English-speaking audience.

*Georg Lind, Hans A. Hartmann, Roland Wakenhut*
Foreword

Lawrence Kohlberg

This volume presents a body of recent research and theorizing done in Germany, providing the English-speaking reader with a very valuable addition to the research on stages of moral judgment and reasoning carried out in America and England. It is unique, and differs from the American work in three regards. The first is in its methodology, which relies on the stage measurement instruments developed by Georg Lind and his colleagues in order to systematically differentiate and integrate the assessment of the cognitive structure or stage of moral judgment and its content, that is, the attitudinal, affective, or normative content of moral judgment. Second, it focuses on the moral development and functioning of both adolescents and adults in a variety of educational and work settings, a topic on which relatively little research has been done in the English-speaking world. Third, it explicitly addresses the socio-political dimensions of moral judgment and functioning.

The basic research instruments of the contributors to this volume are psychological tests and interviews. However, their perspective is primarily socio-logical, directed toward the moral socialization resulting from institutional memberships and the effects of such socialization on socio-political attitudes and behavior.

For those unfamiliar with the theory and research developed in America by myself and my colleagues, I would especially recommend Hans Hartmann’s concluding historical review of philosophic theory as well as Georg Lind’s chapter on psychological theory, which provide excellent introductions to the cognitive developmental theory of moral judgment. Jürgen Habermas’ provocative chapter on the theory is part of an ongoing dialogue between us, raising a number of sympathetic but important and pressing questions about the relation in my theory between moral stages and political ideologies. (Rather than responding here to his questions, I refer the reader to my other responses [Kohlberg, 1984, chs. 3 and 4; 1985]).

Placing this book in context, we may note that the American research on moral stages has centered on a complex structural and hermeneutic analysis of longitudinal data in the U.S., Israel, and Turkey, including interviewing subjects from age 10 through the 30s (the 20s in Turkey and Israel). The method and quantitative data are published in Colby et al. [1987], the theoretical assumptions and interpretations in Kohlberg (1984). Core American research studies are published in Kohlberg and Candee [1984]. As an alternative methodology for assessing moral stages to the qualitative moral judgment interview, James Rest (1979a) developed and validated a multiple choice measure of moral judgment preference, the Defining Issues Test. This test evolved from his dissertation (see Rest, 1973), recently replicated more clearly by L. J. Walker (1983, in press). Rest and Walker have each found that com-
prehension of stages other than one's own forms a hierarchical Guttman scale, with no comprehension of stages more than one stage above one's own and with comprehension of all stages below one's own. Walker has found that subjects prefer the highest stage they comprehend, with no preference for higher stages than the subject can comprehend (unless a subject exhibits a general preference for abstract or flowery language). Based on these findings, Rest developed the Defining Issues Test is based on rating of stage-prototypical statements as most important. This test correlated well with age and education but is not simply a measure of verbal intelligence or achievement, since it correlates with other moral experiences and attitudes when IQ and level of education are controlled.

However, based as they are on my early stage definition and scoring system (Kohlberg, 1958), the Rest statements confound stage structure with a focused content issue or norm; for instance, Stage 4 social system reasoning is confounded with a concern for law and order. In our current scoring system [Colby et al., 1987], the norm of law is found to be used at every stage and clarification by content (law) precedes clarification by stage. In a roughly similar way, the methodology of Lind and his colleagues gets preference scores or content (pro and con), as well as a stage. Since preference is determined by both content and structure, a scoring algorithm can be arrived at for assigning a pure stage structure score for an individual. Some subjects are more consistent in preferring stage structure than content, a factor considered in the tests of Lind and his colleagues. Thus, individual preferences are factored into stage, content, dilemma, and interactive components in the procedure. I believe this to be a highly promising approach, and look forward to future research from Lind and his colleagues. In particular, I look forward to efforts on their parts to establish the correlation or concurrent validity of Lind's measure with our moral judgment interview, since it is the result of longitudinal work with the moral judgment interview which have established the validity of the moral stage construct and its assessment. This has been established [Colby et al., 1987] by the threefold stage criteria of (1) step by step sequence, (2) structural wholeness or cross-dilemma consistency (represented by a single general stage factor), and (3) hierarchical integration established by Walker's study of comprehension and preference.

Turning from this book's methods to its findings, I am impressed by the various chapters showing the relation of moral judgment to the "life-worlds," or situations in the military and in the university. In our American research we also have constructed special dilemma situations directly relevant to "life-worlds" including those of schools and correctional settings (Power et al., in press). Our studies examine the "moral culture" or moral atmosphere of these institutions, assessing the strength (degree of collectiveness and phase) of collective norms as well as their stage. While using interviews and questionnaires, our approach is more ethnographic than those reported in this volume, in that they assess shared norms instead of aggregating individual
scores. They confirm such sensible findings as those reported here, e.g., that an ideologically “democratic” army life-world is reasoned about largely in terms of obedience and discipline. Both the theory presented in this book and many of its findings indicate the serious concern on the part of contemporary German social scholarship and policy to create educational institutions which will produce a “democratic character” and democratic citizens, a concern shared by myself and other American authors [see Power et al., 1989] who are working in the tradition of John Dewey (1916). It is encouraging to read the findings of Döbert and Nunner-Winkler, Portele, Schenk and Bohm, and Lind on university students’ judgments, all of which indicate greater political participation or actions in higher stage subjects as well as more democratic, liberal, flexible, and progressive attitudes and content in this participation. Equally encouraging are Oser and Schläfl’s findings of increased social commitment and democratic attitudes developed through Socratic moral discussion, a procedure roughly similar to that used by my colleagues and myself in the United States.

In sum, this book raises issues of the philosophic justification of the adequacy of the higher moral stages on the one hand, and of democratic or liberal ideologies on the other. It is, therefore, useful not only as a window into contemporary German social thought but also as a contribution in its own right to the ongoing and now international discussion.
1 The Theory of Moral-Cognitive Development

A Socio-Psychological Assessment

Georg Lind

For some decades in psychology, morality has been understood either as the individual's behavior evaluated on the basis of given socio-moral norms, or as that behavior which was determined by morally good motives and affects. Behavioristic psychology has focused primarily on the question of whether individuals comply with given rules of conduct. Affect psychology has regarded this behavior as determined exclusively by inner dispositions, i.e., by motives, drives, or the like, which in their turn have been traced to a wide array of causes, e.g., to human genes, nursing behavior, or environmental pressure. Accordingly, intervention strategies for education and therapy treatments were designed, in the first case, to weaken conditioned links between stimuli and socially disapproved responses, or, in the second case, to lessen “negative” affects (hate, envy, aggression), and to strengthen “positive” affects (love, justice, guilt, shame).

Both the behavioristic and the affect perspectives on moral behavior have recently been challenged by the cognitive-developmental approach, which postulates that moral behavior can only be truly understood and dealt with in addition to norm-following behavior and moral motives, if we examine the cognitive-structural aspect of human behavior. This approach does not lose sight of the fact that human behavior is continually evaluated on the grounds of socio-moral rules, norms, laws, etc., nor does it deny that affective components are involved in every human behavior. But it points out that moral behavior also depends on the individual's ability to see the moral implications of a situation and to organize and consistently apply moral rules and principles to concrete situations. Concrete situations usually imply more than one rule to be observed, and these multiple demands are likely to conflict with another. Moreover, social evaluations of a person's behavior may considerably deviate from one another and may themselves have to be critically evaluated on ethical grounds. To cope with such situations the individual must be endowed not only with moral affects but also with moral judgment competence, i.e., with the ability for reflective thinking and rational discourse. Hence, psychological intervention must also – or even primarily – be concerned with the cognitive aspects of moral behavior as well as with the instilling of moral affects. Cognitive-developmental extensions of these once important models of moral conduct have been suggested by a number of psychologists since the turn of the century, e.g., by Levy-Suhl (1912), Moers (1930), Hetzer (1931), and Piaget (1977/1932). Piaget was one of the first to develop a systematic theory of moral-cognitive development. His research concentrated on children's development of autonomous moral judgment in the ages 5 to 12, particularly in regard
Moral-Cognitive Development

to rules of children's games and their substantial validity. More than twenty years later, Kohlberg took up and considerably furthered this approach to the study of morality. On the basis of longitudinal research of adolescents' and young adults' moral judgment, he suggested an elaborated set of hypotheses about the nature and course of moral-cognitive development (for instance, see Kohlberg, 1958, 1969, 1979, 1984). Especially through Kohlberg's work, the cognitive-developmental theory of moral judgment has attracted much attention from academics as well as from practitioners. In the field of psychology it has stimulated an immense amount of empirical research – not only in the United States and Canada but, since the seventies, also in Europe, particularly in Great Britain (cf. Weinreich-Haste & Locke, 1983) and West Germany (for example, cf. Döbert & Nunner-Winkler, 1975; Habermas, 1976; Portele, 1978; Bertram, 1978; Lind, 1979a; Eckensberger, 1983; Oser, 1984).

In this essay I shall discuss the concepts and assumptions of cognitive-developmental theory as it has been formulated by Kohlberg, and, on the basis of my own research, suggest some modifications and extensions to improve the consistency, the scope, and the empirical validity of the theory. Above all, I shall offer two suggestions: first, that “structural wholeness” is a methodological criterion and not an empirical hypothesis, and second, that Kohlberg's stage model is truly supplementary to Piaget's phases from heteronomy to autonomy and not as a substitute for them. Finally, I shall discuss some implications of cognitive-developmental theory for the relation of individual moral development to the social environment, e.g., for the concept of interaction of person and social environment, the relation of individual moral judgment competence and social position, and the role of selection mechanisms in socio-moral development.

From “External Effect” to “Internal Structure”

Although the study of moral behavior has a long tradition, there is still much debate about its concepts and its methods. Psychological studies of moral behavior are already part of the “empirical study of the soul” (Erfahrungenseelenkunde) of the 18th century and the flourishing of “moral statistics” in the 19th century (Laplace, Ouetlet, Dufau, Drobisch). As early as the first half of this century, a number of psychological experiments were made concerning the conditions for and the development of morality (for an overview see Neumann, 1931; Pittel & Mendelsohn, 1966).

Many of these studies made a careful differentiation between physical behavior considered merely as localized in space and time, social behavior that is evaluated in regard to external standards, and moral behavior considered as possessing socio-psychological meaning. However, in most of these studies a
“behavioristic” point of view prevailed. In these studies – more often in their research methods than in their theoretical premises – the social-evaluative and psychological-cognitive aspects of moral behavior were excluded from consideration. This research perspective is well represented by the studies by Hartshorne, May, and their collaborators under the supervision of the learning theorist Thorndike (cf. Burton, 1978). As do many of their heirs, Hartshorne and May (1928) assessed moral behavior as a physical phenomenon without reference to socio-psychological categories. They argued that

“no progress can be made [. . .] unless the overt act can be observed and, if possible, measured without reference, for the moment, to its motives or its rightness or wrongness.” (p. 10, italics added)

The rationale behind this “pragmatic orientation” (Burton) was that psychology could gain scientific reputation only if it focused its research on purely methodological considerations. Measurement should be as “objective” and free of “subjective” elements as it is, for example, in physics, and this could be achieved only if a physical conceptualization of behavior was adopted. However, this physicalistic behaviorism confused concept with methods: that is – to use the words of Adorno (1980, p. 84) – it turned the objectivity of the behavioral concept into the subjectivity of the research method. In attempting to avoid value judgments, psychologists actually stripped behavior of any socio-psychological meaning. In impinging upon the measurement of the physical aspects of behavior, it failed to assess what it intended to study. The morality or immorality of human acts cannot be adequately described without recourse to socio-psychological properties of behavior, i.e., external social norms or individual motives and thoughts Moral actions, as Blasi (1983) explains, “are responses to situations, as defined by and interpreted according to moral reasoning structures, i.e., to a set of criteria determining the morally good” (p. 196).1

Although today there seems to be little disagreement over the fact that any research of moral behavior needs to take reasoning structures or moral criteria into account, the question remains as to which we should choose “to determine the morally good”. Basically, there are two perspectives of research, the one already mentioned conceives of moral behavior almost exclusively from the social-evaluative perspective, the other from the individual's motives point of view. Allport (1961) has aptly labeled these two perspectives the external effect approach and the internal structure approach, respectively.

The external effect approach categorizes an individual's behavior according to socially given norms, laws, or regulations. This implies that behavior is categorized according to traits which are common to all individuals of a group or sample of persons (common trait approach). Though often the rationale of the
measurement process is not made explicit but is hidden behind the implicit assumptions of the research method, in this approach behavior is judged according to whether it conforms to, or deviates from social norms and expectations, i.e., whether social rules and laws are transgressed or obeyed. In a typical research design the psychologist assesses a subject's behavior according to external social categories like deceiving/not deceiving, stealing/not stealing, killing/not killing. Pittel and Mendelsohn (1966) have found in their review of half a century of psychological research on moral behavior that most methods of assessing moral behavior are indeed based on normative or other evaluative standards of “correctness” determined by societally defined criteria. Thus, responses in agreement with norms established by the investigators are scored as moral, while those not in agreement lower the overall measure of strength of moral attitude or conscience [...]. Even when scoring criteria are not explicitly linked to normative or societal standards, subjective scoring procedures and ratings [...] frequently rely on the same sort of external standards of evaluation. “(p. 33)

The external effect approach to moral psychology can be criticized on several accounts. But, as Pittel and Mendelsohn (1966) noted, “perhaps the greatest single shortcoming underlying each of the specific criticisms discussed is the failure to view evaluative attitudes as subjective phenomena whose measurement is best achieved independent of a concern with the relationship of those attitudes to conventional and normative standards of moral valuation” (p. 34).² Because it is confined to the tacit evaluation of behavior according to socially given norms, the external effect approach fails to assess the cognitive and affective aspects of individual behavior. When studying moral behavior from an external point of view, one must assume that the system of norms is monolithic, and that these norms have an immediate effect upon behavior, i.e., that there is no need to assume mediating processes on the side of the individual. This view overlooks the fact that only when the individual accepts moral principles and orients his behavior to them do these principles become actual. Therefore, we may dismiss the external effect point of view as too narrow. In opposition to the prevailing view, Allport (1961) has already demanded that psychological analysis of the individual personality should focus on the internal structure of human behavior. Similarly, Pittel and Mendelsohn (1966) have called for a change in the approach to moral behavior; for them “it is important to assess at an individual level the content, strength, and patterning of the subjective attitude of evaluation per se” (p. 34). Modern psychology seems to be ready for a psychological interpretation of the term “behavior,” which “includes much of that which in other places is designated as thoughts, feelings, or ideas” (Cohen, 1983, p. 3).

It seems that the cognitive-developmental theory of human behavior has in-
Moral-Cognitive Development

deed succeeded in working out a research program on the basis of the internal structure point of view. The change in the research paradigm is marked by the change of terms from “moral behavior” to “moral judgment” or “moral judgment competence.” But the concept of behavior or performance has not been completely abandoned because, as Habermas has succinctly noted, “competence by itself cannot be shown to exist except in its concrete manifestation, that is, through phenomena of performance” (ch. 1, in this volume). As an empirical science psychology is closely linked to the observation of behavior, because every empirical hypothesis concerning the content, structure, and development of moral reasoning must be verifiable or falsifiable by referring to a manifest pattern of judgment behavior (cf. Kohlberg, 1979a). Nevertheless, we prefer the concept of moral judgment because it refers more emphatically than the term “moral behavior” to the fact that all (physical) human behavior is its topic, but only behavior which can be related to the individuals' own moral categories.

The differences in outcomes of the external effect and the internal structure point of view are marked if one considers, for example, the moral behavior of children. Children are often seen as lacking morality. But this is true only if morality is defined with regard to adults' norms. For example, it is true that “honest” behavior becomes more frequent as children grow older, as does “dishonest behavior in some children (Hartshorne et al., 1930; Block, 1977, p. 40). However, if one considers the reasons for behavior beyond particular social norms and the point to which the cognitive aspect of moral behavior has developed, it becomes clear that even in young children behavior is consistently organized according to rules, although these may be individual, i.e., widely varying and not socially approved. Moreover, moral behavior becomes not only more consistent and integrated with age but also more differentiated. Thus a child who has attained significant autonomy regarding the moral principle “Thou shalt not lie!” will no longer judge a violation of this rule as always wrong. Rather, the child will also consider the circumstances, in fact judging: “It depends.” For example, children who at first consider lying to be generally prohibited, later consider it to be all right if they can keep a friend out of trouble by lying (Bull, 1969, p. 210). Similarly, in the course of the child's cognitive-moral development, the rule of “returning like for like” is differentiated by the idea of mitigating circumstances (Piaget, 1973b, p. 304). For the superficial outside observer, those cases may appear to be morally regressive, whereas a cognitive-developmental psychologist would recognize in them progress in the child's development.

The fact that the terms “moral behavior” and “moral judgment” often indicate two fundamentally different ways of viewing morality in psychology rather than describing two different types of human behavior, renders the study of their interrelation particularly difficult (Blasi, 1980; Eckensberger, 1983; 

[26]
Kohlberg & Candee, 1984). This has been aptly pointed out by Don Locke (1983a, 1983b, 1983c), who described the basic difference between the external and the internal approach as the choice between an “evaluative” and a “neutral” definition of moral action: “We can either define moral action with reference to our own attitudes about what is right or wrong, good or bad [. . .]; or we can define it by reference to the attitudes and beliefs of the agent” (1983a, pp. 112-3). He convincingly demonstrates that the evaluative definition of morality cannot provide a satisfactory basis for psychological research (nor for educational practice). Therefore, he proposes “to work towards a neutral definition, always bearing in mind that moral action in this sense will not necessarily be action that we would ourselves regard as good or bad, right or wrong” (p. 113).

Thus, if we want to understand the individual, two extremes must be avoided: that of defining personality structure solely from the “outside,” and, at the other extreme, giving a purely idiosyncratic definition in which people are considered totally unique and therefore ineligible for comparison. People are bound to society, without which they would become “total abstractions” (Adorno, 1980, p. 197). However, they are also capable of assuming responsibility and exercising critical judgment on the basis of moral principles. We may therefore assume that, beyond the particular characteristics demanded by the situation, there are general personality characteristics on the basis of which people can be meaningfully compared without restriction to conventions. In this interactionist conception, “personality” is neither a purely external nor a purely internal category; rather it is that which is characteristic of the relationship between the individual and the social environment (see below). This relationship is twofold. On the one hand, moral behavior presupposes a cognitive structure: moral principles, norms, and values have to be balanced off against each other and in light of the specific circumstances of a decision situation. On the other hand, competence in moral judgment, i.e., the ability to integrate and differentiate moral principles and apply them to every-day decisions, has a developmental character and so must be placed in reference to the individual’s life experience (ontogenesis) and to the state of the socially developed strategies for solving problems (phylogenesis).

Assessment of the Theory

There is no theory of morality, moral behavior, and moral development that is fully elaborated and empirically proven – as Lakatos (1978) has shown, there can be only preliminary theories which are subject to alteration and error. Nevertheless, we can identify and deal in a critically constructive manner with a number of core concepts and assumptions, which have been empirically well
Moral-Cognitive Development

supported, and which we may provisionally call the theory of moral-cognitive development. This arose out of the tradition of Kant, Baldwin, Janet, and Dewey, and has been elaborated upon extensively by Piaget and expanded and repeatedly varied by Kohlberg, on the basis of numerous creative as well as critical studies.4

Although like Kohlberg we will use the terms “theory” and “approach” interchangeably, we prefer to speak of an approach rather than a theory to make clear that our concern is not only with empirical hypotheses but also with the conceptual framework which provides the “positive heuristics” (Lakatos). The concepts and the hypotheses of any theory have to be analyzed in different ways and must not be confused. Concepts provide the basis for measurement and empirical testing, but they cannot themselves be empirically tested in a meaningful manner. Thus concepts may be assessed as to their usefulness in enlarging our intellectual capacity and discovering new facts and in regard to their internal consistency and absence of contradictions. Hypotheses in which these concepts are used to make statements about (causal) relations in the empirical world can be analyzed in two ways. First, they can be assessed with respect to the information content, i.e., the degree of their unforeseeability and a priori unlikeliness (cf. Popper, 1968; Lind, 1984a). Hypotheses with high a priori probability are – even if empirically true – of little theoretical and practical interest because they are likely to be supported by chance, and provide no information which deserves to be empirically tested and conserved in a scientific theory. Second, informative hypotheses can be assessed in regard to their empirical validity. Critics as well as defenders of cognitive-developmental theory have not always been aware of this distinction, a fact which has caused some confusion in the recent discussion about the value of this theory (cf. Kohlberg, 1976; Phillips & Nicolayev, 1978; Lapsley & Serlin, 1984). In this presentation we want first of all to examine the central concepts and schema of cognitive-developmental theory and then to consider the extent to which the fundamental hypotheses of this approach have proven to be empirically valid – or may have to be treated as concepts which cannot be empirically tested. In particular, we want to make four points. The first is that, as has already been indicated above, the cognitive-developmental theory has provided very useful concepts for the study of moral behavior and thus has enabled us to make new discoveries in the field of moral psychology. The second point is that, in comparison to other psychological approaches, the empirical hypotheses (e.g., about the invariant sequence of moral-cognitive development, preference order of moral stage-types, and moral-cognitive parallelism) are both highly informative and well supported by empirical data. The third point is that the central concept of structural wholeness should not be construed as implying empirical predictions, but should be set up as the criterion against which the theoretical validity of the measurement of moral competence can be
evaluated. Our fourth, and perhaps most important, point is that Kohlberg's stage schema of moral development does not include and replace that of Piaget, but rather succeeds in supplementing and extending it in regard to the social dimension of individual development.\footnote{5}

Conceptual Framework: Structures and Stages

In cognitive-developmental theory, the cognitive-structural aspect of moral judgment and the stage scheme of moral development assume the position of core concepts. Sometimes they have been treated as empirical hypotheses which they certainly are not. Although their meaning has not yet been completely clarified, these concepts obviously cannot be dismissed without giving up the cognitive-developmental approach altogether. When we discuss them, we must keep in mind the fact that both concepts are frequently used in many different ways and have quite different connotations outside this approach (cf. Boesch, 1983; Entwistle, 1979; Glaserfield & Kelley, 1982).

Structures

In cognitive-developmental theory the concept of cognitive structure is usually juxtaposed with the concept of affective content (cf. Kohlberg, 1958, 1969; Lind, 1985a,b). Kohlberg claims that, whereas traditional psychology has focused mainly on the content of moral behavior, his theory is dealing with its structure. It is often considered one of the major tasks, if not achievements, of this approach to distinguish both components and to devise an instrument which makes it possible to measure the structural aspect of moral judgment competence in addition to, and apart from, its content aspect. Although this topic has already been discussed, one can still reasonably ask, “Exactly what is structure and what is content?” (Lickona, 1976, p. 13). Is it the difference between opinions about concrete action dilemmas and the moral reasons given for them? Is it the difference between moral beliefs and moral attitudes? Or does content denote observable behavior, whereas structure is something “behind” behavior and therefore unobservable?

For cognitive-developmental theory the structure of moral judgment behavior reflects the organization and process of moral thinking, the way in which and the degree to which moral maxims or principles are brought to bear in specific situations. The concept of cognitive structure refers to Kant's concept of Urteilskraft (power of judgment) which is required

“partly in order to decide in what cases they (moral maxims) apply and part-
ly to produce for them an access to man's will and an impetus to their practice. For man is affected by so many inclinations that, though he is capable of the idea of a practical pure reason, he is not so easily able to make it concretely effective in the conduct of his life.” (Kant, 1788/1949, p. 52)

This implies that moral action requires both “power of judgment” and real acceptance of moral maxims, i.e., the cognitive ability to understand how a rule is to be applied in concrete situations and the motivation or will to base one's action on rational insights (Habermas, 1983, p. 191). Accordingly, Kohlberg defined moral competence as “the capacity to make decisions and judgments which are moral (that is, based on internal principles) and to act in accordance with such judgments” (Kohlberg, 1964, p. 425); he defined the otherwise neglected cognitive and structural aspect of moral behavior as “the degree to which any of an individual's judgments approximate the criteria of a moral judgment” (Kohlberg, 1958, p. 7). These are the criteria which suffice for the categorical imperative: “impersonality, ideality, universalizability, pre-emptiveness, etc.” (Kohlberg, 1971, p. 215).

Thus, the organization of a person's moral judgment behavior is not characterized solely by the moral norms it serves (or fails to serve), which we may call the affective content of behavior, nor solely by the formal properties of the individual's reasoning, i.e., the consistency or structure of reasoning. It is only by referring to content that one speaks meaningfully of behavioral consistency. There is no consistency of behavior as such; it is always consistency in relation to a criterion or principle. In other words, consistency is a bivalent relation concept. Whereas purely formal structures, as found in physical and chemical nature, are arrangements of elements without dynamic-affective meaning, dynamic structures refer to human actions which possess a meaning, e.g., to behavioral elements which are characterized through a teleological, affective content and their relationship to this and other affective contents (Figure 1). For this reason we cannot define moral behavior without reference to particular moral principles. Yet, following D. Locke's neutral definition of morality, we should refer in psychological assessment rather to the subject's own moral principles than to those of an external judge.
Some psychologists – even cognitive-developmentalists – tend to view these two aspects as separate faculties of the mind, a tendency which is evident not so much in theory as in concrete research methods. If we define the formal structure merely as a “system of inner relationships,” these relationships are themselves purely formal and so lack an affective, dynamic dimension. Formal structures suffice for a mechanistic image of humans taken from association psychology, but a psychological definition of cognitive structures requires a teleological content (cf. Reese & Overton, 1970).

According to our model, affect is provided in a supplementary fashion by the moral contents – moral norms, issues, principles. These moral contents render a behavioral structure “comprehensible” by scientific analysis (Kohlberg, 1979a, p. 14). Moreover, they direct and motivate the development of the cognitive aspect of moral judgment: “The kind of affective pressure which inevitably must operate on the mainly attitudinal premises on which a moral judgment is based are more likely than ever to favor the survival of the existing structure” (Weinreich-Haste, 1975, p. 208). Therefore, Kohlberg's occasional insistence that cognitive-moral stages are, or are to be, defined solely by formal aspects deserves to be criticized. We shall return to this point below. Does this imply, as Glaserfield and Kelley (1982, p. 157) and others have maintained, that structures are merely a heuristic device in the head of the observer but are not themselves observable? I do not think it does. There are examples, even in psychology, of conceptualizing behavioral structure as an observable entity and measuring it directly. Of course, a single act, e.g., expressing an opinion on a particular dilemma, cannot be used as a basis for relational inferences by an observer. At least two acts are needed to suggest a relationship. Moreover, a single act is usually the result of an individual's reaction to the multiple demands of a complex situation. Hence, to make valid inferences on the content and structure of moral behavior, we have to make use of relational information and observe a whole pattern of acts in a particular behavior context and the reasons that justify them. To adequately under-

![Diagram: Figure 1 Purely Formal vs. Dynamic Structures.](image-url)
stand which moral attitudes a person has and to what degree his or her behavior is actually determined by such principles, or, alternatively, to what degree these are used merely to support ("rationalize") unreflected opinions and habits, a hermeneutic circle of hypothesizing and verification is required.

To give an example of such a hermeneutic process, let us consider a discussion about the euthanasia dilemma. A person may say that the doctor who committed mercy killing was morally right. From this single statement we cannot infer whether or not the person based her judgment on moral concern. We must ask for her reasons. In a discussion, she expresses a high acceptance for a Stage 5 argument. Although we now know a little more, we still cannot be sure that she reasons at Stage 5. Her acceptance may be determined by several considerations; the moral quality of the argument is only one among many. She may accept it because the argument is presented by an authority, for example, by a doctor or a psychologist; or she may accept it because the argument supports her independently established opinion on euthanasia. Therefore, we would be more certain if this person did not also accept all other arguments presented in the discussion, and if she accepted the same moral reason even on occasions when it was at variance with her intuitively gained opinion. If this were not the case, we would usually say that her judgment was not determined by the moral principle on which the argument was based. In basically the same manner, this everyday process of probing to advance and eliminate competing explanations for a person’s action is also employed in tests of moral judgment competence (cf. Colby, Kohlberg, et al., 1984; Lind, 1982b; Lind & Wakenhut, in this volume).

We may summarize so far by saying, first of all, that the two aspects of moral behavior, viz. cognition and affect, are both attributes of the same behavior. They can be differentiated only analytically and cannot be regarded as two ontologically separate entities. We shall refer to this basic claim as the Non-Separability Axiom, whose corollary is that the cognitive aspect that refers to the structure of judgment behavior must always be determined in relation to moral content (see also Lind, 1985a). Secondly, both aspects, affective content as well as cognitive structure, are observable in principle; in cases in which this is not possible in practice we have to refrain from an assessment. Thirdly, if acting according to principles involves equilibration, rationality, freedom from contradiction, and suitability, then structural wholeness of reasoning cannot mean a simple consistency or rigidity; concrete behavior which truly reflects the multitude of moral implications of a situation for an individual’s value system necessitates integrated and differential judgments (Kohlberg, 1958, pp. 8-9; 1969, p. 348). Fourthly, the cognitive-developmental concept of cognition is at odds with those treatments which discuss cognition not as a set of general structures and processes, but rather as particular contents, such as beliefs or mental achievements. Fifthly, the notion of structural transformation as distin-
Stages

To provide a conceptual framework for analyzing moral-cognitive development, Kohlberg has constructed six stages, or three levels each of which includes two stages of development. These are well known and shall be summarized only briefly.

Each of the Kohlberg levels of moral-cognitive development is primarily defined by the “socio-moral perspective” which the actor takes in making decisions on socio-moral problems (cf. Kohlberg, 1976). On Level I, the individual assesses a situation from the “concrete individual perspective.” The morally right or wrong is determined by the material consequences of an act; the guiding principle is to avoid punishment and to satisfy one's needs. On Level II the person takes over the “member of society perspective” from which the maintenance of social relations and order becomes an important principle for assessing a dilemma situation. On Level III the actor makes judgments from a “prior to society perspective,” i.e., on the basis of general principles which are not tied to a particular social group or society but to humanity and human life as a whole. These three levels are further subdivided by Kohlberg into two stages, yielding the six stages of moral-cognitive development which are reproduced in Table 1. We have adopted the description of the six stages from Kohlberg and Turiel (1971, pp. 415-6) because these seem to be among the clearest of the many formulations given (cf. Lind, 1976, p. 125; Krämer-Badoni & Wakenhut, 1978a, p. 218; Montada, 1983, p. 7).
Table 1: Stages of Cognitive-Moral Development.

I. PREMORAL LEVEL

Stage 0. Subject neither understands rules nor judges good or bad in terms of rules and authority. Good is what is pleasant or exciting; bad is what is painful or fearful. Has no idea of obligation, should, or have to, even in terms of external authority, but is guided only by can do and want to do.

II. PRECONVENTIONAL LEVEL

At this level the child is responsive to cultural rules and labels of good and bad, right and wrong, but interprets these labels in terms of either the physical or the hedonistic consequences of action (punishment, reward, exchange of favors) or in terms of the physical power of those who enunciate the rules and labels. The level is divided into two stages:

Stage 1. The punishment and obedience orientation. The physical consequences of action determine its goodness or badness, regardless of the human meaning or value of these consequences. Avoidance of punishment and unquestioning deference to power are valued in their own right, not in terms of respect for an underlying moral order supported by punishment and authority (the latter being Stage 4).

Stage 2. The instrumental relativist orientation. Right action consists of that which instrumentally satisfies one's own needs and occasionally the needs of others. Human relationships are viewed in terms similar to those of the market place. Elements of fairness, reciprocity, and equal sharing are present, but they are always interpreted in a physical or pragmatic way. Reciprocity is a matter of “You scratch my back and I'll scratch yours,” not of loyalty, gratitude, or justice.

III. CONVENTIONAL LEVEL

At this level, maintaining the expectations of the individual's family, group, or nation is perceived as valuable in its own right, regardless of immediate and obvious consequences. The attitude is not only one of conformity to personal expectations and social order, but of loyalty, of actively maintaining, supporting, and justifying the order and of identifying with the persons or group involved in it. At this level, there are two stages: [35]

Stage 3. The interpersonal concordance or “good boy – nice girl” orientation.
Good behavior is that which pleases or helps others and is approved by them. There is much conformity to stereotypical images of what the majority perceives as “natural” behavior. Behavior is frequently judged by intention: “He means well” becomes important for the first time. One earns approval by being “nice.”

Stage 4. The law and order orientation. There is orientation toward authority, fixed rules, and the maintenance of the social order. Right behavior consists of doing one's duty, showing respect for authority, and maintaining the given social order for its own sake.

IV. POSTCONVENTIONAL, AUTONOMOUS, OR PRINCIPLED LEVEL

At this level, there is a clear effort to define moral values and principles which have validity and application apart from the authority of the groups or persons holding these principles and apart from the individual's own identification with these groups.

Stage 5. The social-contract legalistic orientation. Right action tends to be defined in terms of general individual rights and of utilitarian standards which have been critically examined and agreed upon by the whole society. There is a clear awareness of the relativism of personal values and opinions and a corresponding emphasis upon procedural rules for reaching consensus. Aside from what is constitutionally and democratically agreed upon, the right is a matter of personal values and opinion. The result is an emphasis upon the legal point of view, but with further emphasis upon the possibility of changing the law by appeal to rational considerations of social utility, (rather than on rigidly maintaining it in terms of Stage 4 law and order). Outside the legal realm, free agreement and contract are the binding elements of obligation. This is the acknowledged morality of democratic government and constitution.

Stage 6. The universal ethical principle orientation. Right is defined by the decision of conscience in accord with self-chosen ethical principles appealing to logical comprehensiveness, universality, and consistency. These principles are abstract and ethical (the Golden Rule, the categorical imperative) and are not concrete moral rules like the Ten Commandments. At heart, these are universal principles of justice, of the reciprocity and equality of human rights, and of respect for the dignity of human beings as individual persons.

Despite numerous variations in publications by Kohlberg and others over the last two decades, the stage schema has remained the same in its essential components (cf. Bergling, 1981; Kohlberg, 1983). Among the changes was the in-
Moral-Cognitive Development

introduction of a so-called Stage 4 1/2, so as to incorporate unexpected regressions in passing from Stage 4 to Stage 5 (this modification was subsequently abandoned in large part; today Kohlberg regards this problem as a coding error), and the attempt to expand the model to include a seventh stage (Kohlberg, 1973; Habermas, 1976), which was set aside when Kohlberg, in response to academic criticism, omitted Stage 6 from his research program (but cf. Kohlberg, [1986], for his subsequent rethinking of Stage 6). In addition to these changes in the basic stage concept, the problem of assigning persons to stages on the basis of complex responses, the need for a more finely graded scale of moral-cognitive development, and the problem of relating moral thought to action have led Kohlberg and his associates to add a number of mixed stages and substages whose theoretical meaning is mostly unclear and which seem to contradict the central assumption of “structural wholeness” (which is true if this assumption is understood to be an empirical hypothesis; I will return to this point below). I believe that the differentiation into A and B substages is one of the most important changes in the stage model (Kohlberg, 1978, p. 71; for an extensive account see Kohlberg, 1984).

However, this differentiation necessitates a revision of our view of the relation between the models proposed by Kohlberg and Piaget. Kohlberg had postulated that his schema of moral-cognitive development substitutes for Piaget's model and extends it upwards (Kohlberg, 1958, pp. 151-228; Weinreich-Haste, 1975, p. 206), by which he means that Piaget's phases of heteronomy and autonomy are equivalent to his Stages 1 and 2 respectively. Stages 3 to 6, he supposed, extend beyond Piaget's developmental scale. This position can be justified only if (a) it is true that the moral judgment competency of a person manifests itself simultaneously in all areas of life, and (b) the age grouping in Piaget's two phases and Kohlberg's Stages 1 and 2 are indeed the same (cf. Kohlberg, 1958, pp. 70, 377-383; Kohlberg 1969; Bergling, 1981).

But both assumptions are questionable. First, “structural whole” does not in the least have to mean that moral autonomy is acquired all at once in all areas of life in relation to all moral issues. Even if his notion of “horizontal décalage” seems to imply the opposite, Piaget stated clearly that “there are no global stages that would characterize the complete psychological life of a subject at a particular time in development” (1973, p. 91). Second, I do not believe that it is permissible to set Piaget's phase of moral autonomy as equivalent to Kohlberg's Stages 2 (“hedonistic instrumentalism”) or 3 (“group conformity”). We need not invoke Piaget's concept of logical development to see that the phase of moral autonomy more closely resembles Kohlberg's stage of principled morality and postconventionalism, even though they appear in completely different age groups because of the different issues and norm contents involved. Kohlberg seems to have been aware of this ambiguity when he once equated the phase of moral autonomy with both Stage 2 and Stage 6 (Kohl-
Therefore, I suspect that, as Weinreich-Haste (1975) has suggested, the two concepts do not substitute for but instead supplement each other. Exactly how they relate to one another may become clear if one analyzes the way in which the affective and the cognitive aspects of moral judgment are connected.

Cognitive and Affective Aspects of Moral Stages

In his dissertation, Kohlberg considered two dimensions of moral-cognitive development, the cognitive-structural dimension represented by the three levels, and the affective dimension represented by the six stages. In fact, in the newer versions of his model, he still defines the “structural” stages rather by moral content than by formal categories (cf. Kohlberg, 1976; also Table 1, above). Even abstract moral principles such as equality, justice, and universality are not in themselves structural but are contents if they are conceived of merely as deontic principles to which a subject refers in his or her argument. These criteria can be called “formal” only if the subject’s judgment indeed matches the principles of justice and universalizability. This fine, yet important, distinction seems to be taken up by Kohlberg through the introduction of A- and B-substages. The B-substages assume characteristics very similar to the main Stages 5 and 6, the level of principled morality and moral autonomy, and both criteria are indeed highly correlated in the examples provided by the Kohlberg interview manual (Rest, 1979a, p. 43; Eckensberger, 1984; Lind, 1985a). Moreover, the correlation shows also in the Milgram experiment on obedience, in which for 6 out of the 8 subjects who could be unambiguously assigned to stages, the two classifications coincided (Kohlberg & Candee, 1984, p. 69). Interestingly, this study also shows that the $A/B$ distinction helps us better than does the stage distinction to understand the behavior of the subjects in Milgram’s experiment on obedience. Whereas only two out of 7 substage $B$ persons executed the order to torture another person in spite of his screams of pain (which they did not know were faked), all 9 subject on substage $A$ obeyed this inhuman order.

In accordance with our two-aspect model of moral behavior, and to account for the inadequacies of Kohlberg’s stage model, I have suggested two distinguishable, albeit not ontologically separable, dimensions of moral-cognitive development (Lind, 1978b, 1984a; see also ch. 8, in this volume). Accordingly, moral-cognitive development may be understood as a two-dimensional process, in which Piaget’s phases describe a recurring sequence of cognitive transformations on each of Kohlberg’s stages. Whereas Piaget focused on conflicting norms of children’s games, Kohlberg concentrates on norms of secondary groups and on universal moral principles and values of human life,
replacing the child society in Piaget's work with the adult society. Hence, with his six types of moral issues, which are related to social institutions such as market rules, family and friendship norms, law, and moral principles, a completely new dimension of development has emerged: the differentiation of the affective aspect of moral judgment according to six types of socio-moral perspectives (see Figure 1).

In his recent writings, Kohlberg acknowledges the A- and B-substages as theoretically equivalent to Piaget's phases of heteronomy and autonomy (Kohlberg, 1984, pp. 652-683), though he interprets these substages in a somewhat different manner than I do. When he says that this distinction lies “midway between form and content” (Kohlberg & Candee, 1983, p. 44), he apparently views it as a new entity which is separable from both moral content and structure. This multiplication of psychological entities is, in my view, not necessary, and could be taken as a degenerating problem shift. As I have noted above, cognition and affect are distinguishable aspects of moral behavior but are not separable. The theory of an integral moral-cognitive development is distinct from approaches which either presuppose that the two developmental dimensions cannot be distinguished at all, or attempt to conceptualize two or more ontologically separate components or factors of development.6

Empirical Validity and Information Value of the Theory

The fruitfulness of Kohlberg's cognitive-developmental theory is documented through the wide array of new empirical hypotheses which could be formulated within this conceptual framework. Four hypotheses are especially noteworthy here, concerning as they do (a) the invariant succession of the developmental stages, (b) the structural whole, or organization, of moral judgment, (c) the hierarchical order or preference order of moral reasoning types, and (d) the parallelism between the development of the cognitive and the affective aspects of moral judgment. In the following I want to analyze what these hypotheses imply, whether they can be confirmed empirically, and to what extent they have been empirically validated.
The most central hypothesis of cognitive-developmental theory is that there are qualitatively different stages of moral development which form an invariant sequence; social factors can accelerate or slow down the development but cannot change its sequential order (Kohlberg, 1969, p. 352; 1971, p. 181; Colby et al., 1983, p. 1). This implies that “a single case of longitudinal inversion of sequence disproves the stage theory, if it is not a manifest case of measurement error” (Kohlberg, 1973, p. 182).

This hypothesis is highly informative or testable, a fact which is not adequately reflected in the usually reported percentages, correlations, and tests of statistical significance, because these figures do not take into account the Gehalt (information content) of a hypothesis (cf. Popper, 1968; Meehl, 1978). A high Gehalt (G) means that there is a high a priori probability that the hypothesis is not confirmed by pure chance. For example, the a priori probability that an individual person will pass through the six stages in exactly the sequence prescribed by the theory is only $p = 0.0014$ and thus the Gehalt and information content of such a hypothesis is $G = 1 - p = 1 - 0.0014 = 0.9986$ (maximum: 1.0).

The Gehalt of the hypothesis that, for example, 50 persons will develop as the theory states thus is extremely high, i.e., it is very close to 1: $G_{50} = 1 - 0.0014^{50}$ (power of fifty).

In comparison, the Gehalt of a vague but very common hypothesis like “variable $x$ somehow influences variable $y$” is close to zero. For further details on this method of evaluating scientific hypotheses, and for an exact definition of $p$ and $G$, see Lind (1984a). A similar method has recently been suggest by Turner (1980). However, his index is based on a mixture of a priori probability and empirical frequency. It is important to note that $G$ is based purely on a priori or logical probability implied by the hypothesis.

Although the hypothesis of invariant moral-cognitive development is so highly informative and easily falsified by empirical data, it has by and large been corroborated by longitudinal studies. Infrequent deviations from this hypothesis were mostly found in short-term studies and in studies using spe-
Moral-Cognitive Development

cial kinds of measurement. In the most important study by Kohlberg and his collaborators, which lasted more than twenty years, the developmental sequences anticipated have occurred with only a very few exceptions. In the 58 persons interviewed every three to four years, only 14 out of 193 passages (7%) were reversed (when measured on a scale divided into 13 interval sub-stages; cf. Kohlberg, 1979a; Colby et al., 1983). The invariant sequence hypothesis could also be supported in studies measuring somewhat different aspects of moral judgment. In studies with the Defining Issues Test, it has been shown that persons between 13 and 22 years of age the consistency with which post-conventional arguments are preferred to other kinds increases considerably (about 20 percentage points; see Rest, 1979a, p. 140). In our ongoing longitudinal study we found that the consistency of evaluating arguments with regard to moral principles grows with increasing age and educational experience, and the tendency to rationalize one's opinion by reference to moral reasons decreases (cf. Lind, 1984a, 1985b).

The small number of regressions, as well as the fact that only a few Stage 5 subjects and virtually no Stage 6 reasoners could be found, has stirred up a debate over how these anomalies should be dealt with. Following simple falsificationism, some tend to regard cognitive-developmental theory as falsified by these anomalies and thus argue for an enlargement or fundamental change of the theory. But following Lakatos (1978), I like to search first for methodological imperfections which can be remedied on the basis of an unchanged theory before inventing auxiliary hypotheses and thus reducing the Gehalt of cognitive-developmental theory. Besides we should not give up a good theory before we have a better one at hand (cf. Lakatos, 1978; Kohlberg, 1979a; Lapsley & Serlin, 1984).

The observed cases of regression may indeed be cases of measurement error, which, however, can be determined in different ways. A classical psychometric way is to select a particular sample of people and assess their stage scores twice within a time span to calculate the correlation between these assessments (the so-called test-retest reliability), which provides an estimate for short-term variations (standard error of measurement). On the basis of this criterion, Colby et al. (1983) could show that in their longitudinal sample the number of downward movements over a three-to-four-year interval (approx. 7%) was clearly lower than the number of changes within one month. This is undoubtedly an impressive result. Nevertheless, I would like to question the adequacy of such psychometric criteria. The size of the standard error of measurement depends very much on the distribution of the stages in the sample (Colby et al., 1983, p. 26, report the standard deviation was here as small as 7/10 of a stage), and from a structural point of view one may rightly question the basic assumptions of classical psychometric theory (cf. Kohlberg, 1976, 1979a; Lind, 1982b).
For this reason, we have pursued another way of determining possible sources of measurement error in moral judgment research. Kohlberg and his colleagues have taken great pains to improve the scoring method, but little attention has yet been paid to the design of the instrument itself. There are in particular three indications that the design of the assessment method could indeed be improved in regard to its theoretical validity. First, if one reviews the original Kohlberg and Kramer (1969) cases of regression, one finds that these cases are almost exclusively confined to the initially higher stage subjects, which indicates, as already noted above, that there may be a deficiency of the method in dealing with high stage reasoning. Second, as Broughton has found in an unpublished analysis of a severe case of regression, some regressions may reflect a lack of probing in the interview. Third, if one analyzes not only the number of times in which a subject reasons on each stage but also which (relative) weights he or she attaches to these reasons, as is done, for example, through the *Moralisches Urteil Test* (cf. Lind & Wakenhut, in this volume), then we find that persons with an initially high judgment competence accept Stage 5 and 6 reasoning slightly less after one year, but this decrease is relative only to the person’s own initial acceptance; his absolute preference for high stages of moral reasoning remains much higher than that of the subjects with lower judgment competence (cf. Lind, 1983a); so here again the regression phenomenon seems to be caused by the developmental restrictions imposed upon the data by the research instrument.

As far as the lack of Stage 6 moral judgments and the infrequency of Stage 5 judgments in the research data is concerned, similar methodological considerations apply. To a certain extent, our present methods seem to be biased against moral reasoning at postconventional stages. This may be due to a lack of probing, to the computation of average stage scores (favoring the scoring on Stages 3 and 4), or to the kind of moral issues involved in the research instruments. In sum, our analysis shows that we can consider the invariant sequence hypothesis as empirically valid. Yes, there are some data which are apparently at variance with this statement, but there is no compelling reason to give up the core hypothesis of cognitive-developmental theory or to narrow its range of applicability as some authors have suggested (e.g., Bergling, 1981; Gibbs, 1977; Kohlberg & Kramer, 1969; Kohlberg et al., 1983; Rest, 1979a). Such alterations may considerably diminish the *Gehalt* (information value) of the theory and should be undertaken only after the above mentioned possibilities of methodological improvement have proven to unsuccessful.
Cognitive-developmental theory states that each of the stages of moral judgment forms a structural whole that unifies an individual's judgment behavior. This assumption is usually taken to imply a consistency of answers over different aspects. Moral orientations should appear “as a logical and empirically related cluster of responses in development” (Kohlberg, 1969, p. 353). As a confirmation of this hypothesis, Kohlberg points to the fact that a person uses moral principles largely independently of the specific dilemma, and that, in regard to this, the differences between persons are highly stable.7

Although in most studies considerable variation in individual judgment – that is moderate consistency – has been found,9 this finding could be greatly improved by introducing new concepts and structural criteria. “The basic developmental concept underlying the revised stage sequence is the level of socio-moral perspective, the characteristic point of view from which the individual formulates moral judgments” (Colby et al., 1983, p. 6). This means that one cannot expect all judgments relating to moral dilemmas to be alike but that, if one follows an individual's reasoning to its roots, one will eventually find that he or she argues from a unique socio-moral perspective which is characteristic of his or her moral-cognitive development.

However, this finding does not yet fully support the original claim of the cognitive developmental theory. The fact of mixed stages, of substages, and of the grading of stages into one hundred developmental points involved in Kohlberg's assessment method still contradicts the proposition of whole stages of judgment. The introduction of A- and B-substages partly accounts for this. Finally, the interpretation of structural wholeness as response consistency is too narrow. It disregards differentiation, which is also an important outcome of moral-cognitive development. To overcome this problem, it has been suggested that the six-stage model be replaced by a “more complex stage model” which would do more justice to a specific data state. The “complexity” hypothesis, however, is too imprecise and has little Gehalt, since it cannot be disproved empirically.

Considering these problems, I suggest that this assumption involves – as Austin has called it – a descriptive error. Structural wholeness as the description of a state is misunderstood; it has to be understood as a norm for an approach to the subject and as a norm for the methods used in dealing with the subject (for a discussion of this problem in respect to psychology in general, see Hartnack, 1962, p. 91). If we regard structural wholeness as a methodological criterion of cognitive developmental theory, the degree and kind of consistency with which a person brings a moral rule to bear in his interaction with social situations will gain the status of manifestations of judgment competence (Beilin, 1971, p. 173; Lind & Wakenhut, in this volume). As Piaget
and Kohlberg have noted, structures have not always been present in the individual and do not emerge all at once, but are constructed through the individual's interaction with his or her social environment. At this point a problem may arise because of the right interpretation of the terms “integration” and “differentiation.” If, within the trait model of personality, we translate integration with increasing and differentiation with decreasing consistency of judgment, two mutually exclusive assumptions result, namely that in the course of development, judgment becomes consistent and inconsistent at the same time. This contradiction is resolved only when “consistency of judgment” is defined explicitly in relation to the orientation to which judgment is consistent or inconsistent. If one analyzes the context that defines the consistency of judgment, then it appears that the consistency decreases in relation to argument conformity so that judgment is differentiated. In contrast, consistency increases in relation to the moral quality of the arguments (the “Stage factor”) so that it is simultaneously integrated. Earlier orientations are seldom abandoned; rather they are differentiated according to a new priority rule. It all depends on not losing the ability to make decisions while developing the capacity for moral reflection. One continues to form opinions about concrete moral dilemmas, but these opinions are reflective commitments, which are open to modification through arguments.

Thus both must be true for a moral judgment to be called mature: it must be made on the basis of universally valid moral principles (integrated judgment) and, at the same time, must attend to the particularities of the circumstances of each dilemma and to their specific moral implications (differentiated judgment).

Preference Order

The stages of moral development are not ordered only on the basis of their philosophical adequacy; in fact people intuitively prefer them in this order (Kohlberg, 1969). This hypothesis has been unanimously supported by a high number of studies in various cultural contexts. Because of this the hypothesis of a universally valid order of preferences for the stages of moral reasoning may seem to be a trivial one, but it is not. First, I think this coincidence of philosophical reasoning and “every-day” moral philosophy is most remarkable. If empirically warranted, it would provide the best and possibly the only way to enter a moral discourse with another person which is especially important for parents and teachers who are concerned with moral education. Second, this assumption, like the first hypothesis, has a comparatively high information content. If the order of preference is determined randomly, the six types of moral rules may be ordered in 720 different ways. Hence the information content
of the prediction that a person or a group of persons will prefer them in the theoretically expected order is as high as \( G = 1 - 0.0014 = 0.9986 \). Third, the hypothesis of preference order may explain why the invariant stage-wise development of moral competence is found universally. The affective component of moral judgment may be considered a pacer for the development of the cognitive aspect of judgment: “The disposition to prefer a solution of a problem at the higher level available to the individual partially accounts for the consistency postulated as our third (structural whole) criterion” (Kohlberg 1969, p. 353). It has been found in many studies that the preference of the morally highest stages is indeed much sooner developed than the ability to use these stages in an everyday argument in a consistent and differentiated manner (Rest, 1979; see also Lind, 1985d).

**Moral-Cognitive Parallelism**

Now that we have identified the cognitive and affective aspects of moral judgment, we must ask how closely the two aspects are related. Cognitive-developmental theory hypothesizes that “affective development and functions and cognitive functioning are not distinct realms. ‘Affective’ and ‘cognitive’ development are parallel” (Kohlberg, 1969, p. 349). This hypothesis is at the heart of cognitive-developmental theory, though still not at the heart of research practice (Lind, 1985a). Its meaning has remained rather obscure, in spite of its very different interpretations. The interpretation most important for the theory of cognitive development goes back to Kohlberg's postulate that “a moral act or attitude cannot be defined either by purely ‘cognitive’ or by purely motivational criteria” (1958, p. 16; our italics). In a similar vein, Piaget had already stated that “every form of behavior has an energy or affective aspect and a structural or cognitive aspect” (1976, pp. 7-8; our italics). Because this hypothesis had not been dealt with adequately in the design of research methods, it had not yet been submitted to empirical investigation (disregarding attempts in which the cognitive aspect has been operationalized as a separate mental faculty). To render it possible to test empirically the hypothesis of parallelism, a new research design was needed. We have suggested such a design with the *Moralisches Urteil Test* (MUT; cf. Lind & Wakenhut, in this volume). Through measuring simultaneously the affective and the cognitive functions as aspects of a particular pattern of judgment behavior, we are now in a position to test the hypothesis of affective-cognitive parallelism directly and without a vicious circle. Indeed, all studies with the experimentally designed MUT have shown a clear parallelism between the affective and the cognitive aspects, i.e., between the content and the structure of moral judgment. The greater consistency is in relation to moral categories, the
stronger the acceptance of the “higher” stages of moral argumentation and the rejection of the “lower” stages. In all studies, the pattern of correlations between the two aspects is surprisingly consistent with the theory of moral-cognitive development (Figure 3; see also Lind, 1985e).

Moral Judgment and Social Context

Turning to educational practice, we see that the theory of cognitive development has been increasingly confronted with the problem of the relationship between individual moral judgment and the social environment. As we have seen, this problem has always assumed a prominent place in the work of Piaget and Kohlberg. However, as Bertram (1980) has noted, the working-out of precise hypotheses and the incorporation of them into practical research plans has been incomplete and unsystematic. One exception is the domain of pedagogical intervention, which has already been frequently discussed (see Higgins, 1980; Lemming, 1981; Scharf, 1978; Oser, 1981b). But until now other important fields have been largely neglected, for example the relationship between the individual and the environment from the angle of the interaction of the person with the environment, the cognitive-moral import of positional differentiation in social institutions, and the role that social selection plays in moral development.
Moral-Cognitive Development

Interaction of Person and Social Environment

Hartshorne and May's *Studies in the Nature of Character* (1928-30) is still the best-known attempt to settle the controversy between personalism and situationalism. These broadly planned studies were supposed to answer the question of whether people are really guided in their behavior by stable character traits or whether situational factors can be held responsible for “immoral” behavior, such as lies or deceit. In order to deal with the character trait of “honesty,” children were brought into “natural” performance situations and their behavior was observed.

What were the results of these studies? There were few children who were honest or dishonest in all classes of situations, and few situations in which all children reacted in the same way. From this the authors of the studies concluded that the position of personalism is untenable, because human behavior is not determined by inner motives or attitudes but rather by the specific situation in which it appears. This conclusion has evoked a lively and long-lasting de-
bate (see e.g., Allport, 1929-30; Asch, 1952; Bem & Allen, 1974). From an interactionist point of view, the results actually call both positions into question.

It has been attempted post hoc to save the (external) trait model of personalism by the hypothesis of measurement error. Thus, Burton (1963) tried, adhering entirely to psychological conventions, to explain the deviations by arguing that some of the situations in which the children were studied had led to “unreliable” measurements and that these situations ought in consequence to be eliminated from the analysis. But little is gained by this argument, since the alternative approach of situationism can also be verified by such a post hoc “explanation.” The measurement error hypothesis can also be evoked for this position, if, following Burton’s own procedure, one eliminates the “unreliable” persons from the analysis, i.e., if one eliminates those cases which call the situationist approach into question and thus confirm this position “empirically.”

Schematically simplified, the results of the Hartshorne and May study indicate the following pattern. In several situations there are no differences in moral behavior among the persons (in the scheme below: Situation A and B), whereas in others the person’s behavior is differentiated (Situation C and D). Some children are honest (or dishonest) in all situations (Person 1 and 2), others vary in this respect from situation to situation (Person 3 and 4). Thus, without Situations A and B – between which there is no correlation (due to the lack of variance) and which, by convention, are labeled “unreliable” – there remains in the findings a pattern that confirms the position of personalism.

<table>
<thead>
<tr>
<th>Situations</th>
<th>Persons</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>Behavior:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h = honest</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>d = dishonest</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>h</td>
<td>h</td>
<td>d</td>
<td>d</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>h</td>
<td>h</td>
<td>d</td>
<td>d</td>
<td></td>
</tr>
</tbody>
</table>

Without Persons 1 and 2, between whom, likewise, no correlation exists, there remains in the results a pattern that confirms the opposite position of situationism.

Thus, a purely statistical treatment can support either position and, therefore, is unsatisfactory (cf. Olweus, 1976). The number of correlations and the proportion of variance accounted for by the person, the situation, or the interaction of both in a statistical sense cannot be a criterion for the model with which we try to understand moral behavior. The question is whether person and situation lend themselves at all to being contrasted in this way. In beha-
Moral-Cognitive Development

Behavior, obviously, both person and situation are always involved. The individual and his environment can be clearly distinguished, but neither can be conceived of without reference to the other. The German term for behavior, *Verhalten*, indicates this, as it refers intrinsically to behavior which is not a solitary event but part of the relation (*Verhältnis*) of persons to their environment.

Accordingly, from the viewpoint of cognitive interactionism, as Kohlberg (1971b) has succinctly pointed out, “moral judgments and norms are to be understood ultimately as universal constructions of human actors...rather than as passive reflections of either external facts...or internal emotions” (p. 184).

Positional Fragmentation

The universality of the stages of sociomoral perspective may be due in part to transcultural communalities of social institutions and positions therein. “All societies have many of the same basic institutions of family, economy, social stratification, law, and government. In spite of great diversity in the detailed definition of these institutions, they retain certain transcultural functional meanings” (Kohlberg, 1969, p. 397). The institutions of society furnish definitions of situation that constitute the life-world of the individual (cf. Krämer-Badoni & Wakenhut, in this volume). Since, however, social institutions are finite, historical, and thus incomplete, they intrude at the time of moral development as requesting and supporting obstacles. Freud calls them the third element, the source of social suffering from which the “discontents of civilization” are derived. Obviously, it is possible and even probable that the real moral atmosphere of a social institution will deviate from the idea upon which it is based (Piaget, 1973). Consequently some authors de-emphasize the educational significance of the ideas on which such institutions are based, or dismiss them as “unreal” ideology. Nevertheless, the moral ideology of social institutions is real and, therefore, can provide an opportunity for criticism and renewal of these institutions in that it makes possible an appeal to moral ideas even if, or just because, they are not fully realized in institutions (cf. Habermas, 1976a).

In modern bureaucratic systems, the relationship between the individual and society is characterized by social positions and memberships in social institutions. Max Weber assumes that the understanding of society by its members is “fragmented” accordingly (1968, pp. 472-473). Weber distinguishes between the understanding of a “client,” who regularly obeys because of rewards and punishments; that of a “profiteer,” who conforms by being reliable in the eyes of his superiors in order to gain recognition and social advancement; that of an “administrator,” who enacts procedural rules and regulations in order to insure institutional stability and smooth operation; and, finally, that of an “initiator,”...
who chooses ethical principles, goals, or values freely so as to apply them universally and consistently in existing and future social institutions. The concept of positional fragmentation implies a developmental sequence analogous to cognitive-developmental theory. The individual's sociomoral perspective may be seen to develop through the perspectives of the client (Kohlberg's Stage 1-2), the profiteer (Stage 3-4), the administrator (Stage 5), and the initiator (Stage 6). Note that in this developmental scheme the phases of “institutionalization” and “complete approval of collective norms” are understood as leading to greater moral autonomy on the part of given institutional groups and thus are closer to Piaget's view than to Durkheim's. For a further discussion of the relationship between individuals and institutions see also Lavoi & Culbert (1978), Spence (1981), as well as several chapters in this volume.

Social Selection

A number of results from research in moral psychology allow the interpretation that the development of moral judgment is also bound up with processes of social selection, and that the environment determines through selection which forms of judgment are to be found within it. Kohlberg (1958) showed that children with a higher level of moral judgment were preferred by their friends (p. 75). Findings such as these are considered, for the most part, as proof of the causal significance of social participation in the process of moral development. This relationship, however, can also indicate the opposite causal relationship, i.e., that the possibility for participation depends on the moral development. If this holds true, the selection (other or self) organized according to the individual state of development has far-reaching consequences for the constitution of society and changes within it.

Selection processes apparently play a remarkable role – be it a positive or negative one – in individual as well as social development. Studies have shown that the state of a child's moral development has little influence on mere participation in social activities. But its significance is related to social prestige in the eyes of classmates and teachers (Keasey, 1971). Furthermore, teachers are quite capable of estimating the moral level of their pupils (Kohlberg, 1969, p. 394; Peck & Havighurst, 1962, p. 183). What consequences this has for the person upon whom judgment has been passed is difficult to tell. Some studies show that children tend to choose as their leaders persons showing a comparatively high moral stage (Keasey, 1971; Peck & Havighurst, 1962, p. 198). This finding coincides with the idea that “leader positions require [...] 'rules and justice' forms of role-taking” (Kohlberg, 1969, p. 399). The higher the social position of an individual, the more he is objectively responsible for decisions in society and “the more he must take the roles of others
Social selection along moral lines seems also to be involved in admission policy at the level of large social institutions. Portele (in this volume) has found that there is a close relationship between the standardization of fields of study and the moral consciousness of academics. In our studies we found that this relation is not due only to university socialization; these differences exist in part even before people enter the university. Likewise, the differences in moral judgment competence among soldiers, officers, and conscientious objectors seem to be linked only partly to the effects of their social environment. In part the differences exist even before they enter into these environments (Lind, 1984a; see also Lippert, 1981). In any case, important questions are hinted at that concern the working together of socialization and selection processes.

**Conclusions**

After enjoying a euphoric and uncritical initial reception, the theory of cognitive development, as formulated by Jean Piaget and Lawrence Kohlberg, has in recent years come up against opposition of an often undifferentiated and irrational nature. Especially in Kohlberg's work—typical of new, creative research paradigms—there are more than a few contradictions and inaccuracies that offer critics abundant points of attack.

The analysis presented here of the basic assumptions in cognitive-developmental theory and of the current findings to date show that we are concerned with an approach which should, in fact, be taken seriously. In several respects the tendency to immunize the theory has been rightly criticized. We do not, however, see any scientific reason that would justify referring to the cognitive-developmental approach as a "degenerating research program" or as a sterile approach. On the contrary, as the body of accumulated empirical research shows, the cognitive-developmental approach has very informative hypotheses at its center that are— with some exceptions— verifiable and verified. Therefore, we should regard the cognitive-developmental approach on the whole as a very "courageous speculation" which has proved to be of great significance for progress in moral psychology, even in the areas where it fails.

The most important innovation of the cognitive-developmental theory, in my view, is of a conceptual nature. It renders the concept of behavior more psychological by recourse to its affective and cognitive qualities, and it renders the cognitive aspect of moral judgment assessable in practice. I view as its core assumption a cognitive-affective parallelism in the development of moral thought and action, which presupposes a two-dimensional model of development. The differentiation of the developmental model into two dimensions or aspects should, however, be strictly distinguished from a bifurcation of cogni-
tive-moral development, in which both aspects are conceived of (and operationalized) as separate faculties of mind. Moral content and moral structure are not composed of insulated behavioral acts which are accessible in an isolated state. Instead they are, as the concept of structural wholeness indicates, characteristics of a behavioral totality and thus have to be dealt with as an inseparable entity.

The methodological and pedagogical consequences of this theory have hitherto not received the attention they deserve. Two such consequences seem especially noteworthy. First, in the field of moral psychology we have to think about a new psychometrics which takes into account simultaneously both the affective and the cognitive aspects of behavior rather than interpreting cognitive characteristics as “errors” of the measurement instrument. We have dealt with this elsewhere in this volume, as well as in Lind (1982b; 1985a). Second, in the field of education, the distinction of the two aspects may help us better understand the cognitive nature of within-stage development and the best ways to foster this. Although modern pedagogy wants to refrain from the indoctrination of moral contents, it is seen to be responsible for stimulating the cognitive aspect of moral growth, i.e., the development of integrated and differentiated judgment.
Notes

1. See also Moers (1930) who stated that behavior “becomes good or had only through its motivation” (p. 441), because “the act that is without real insight and conforms to ethical norms only because of chance events in one’s education or adaptation is not yet a truly good act” (p. 440). Similarly, Hartshorne and May (1928) postulated that “the essence of an act is its pretense” (p. 101), though this remained a play on words which had no real consequences for their research methods.

2. Piaget is exempted from this critique by Pittel and Mendelsohn (1966). Kohlberg's work was not included.

3. See Habermas, 1976a; Schluchter, 1979; Hartmann, in this volume.

4. For an overview and critical evaluation of recent moral judgment research we refer primarily to Bergling, 1981; Bertram, 1980; Blasi, 1983; Broughton, 1978; Colby et al., 1983; Eckensberger, 1983; Habermas, in this volume; Kohlberg et al., 1983; Lempert, 1982; Portele, 1978; Rest, 1979a.

5. Piaget's approach to the relationship of morality and environment has been revived by Bertram, in this volume, and Oser, 1981b.

6. See, above all, the complex “spiral model” by Eckensberger (1984; also Eckensberger & Reinshagen, 1980), the “two-factor model” by Nisan (1984), and the two-component model by Lempert (1982). I cannot discuss these sophisticated models as extensively as they deserve, but I should mention my concern that they also tend to multiply entities and thus to view content and structure-affect and cognition-as separate things rather than as two aspects of one and the same behavior (see also Lind, 1985e).

7. “Factor analysis indicates a single ‘stage’ factor cutting across all moral situations and all aspects of morality on which the individual is assessed” (Kohlberg, 1971b, p. 177). See also Kohlberg, 1958, pp. 11, 338; 1969, pp. 368, 389; 1976, p. 47; 1979, p. 21; Rest, 1979a, pp. 50-51.


10. See also Kohlberg, 1969, p. 434; 1971b, p. 186; Piaget, 1977; Lind, 1985e.
Three Theoretical Approaches

Cognitive-developmental theory has been credited with a notable contribution to the study of the development of the democratic personality. In a review of “political socialization and models of moral development,” Friedman (1977, p. 361) asserts that “the Kohlberg model of man, i.e. the last stages toward which his system is directed, is a model which is democratic in process and goal.” In this article we shall outline this conception together with two other “models of man” to demonstrate that, in spite of its great utility, the cognitive-developmental theory and methodology may be improved upon by taking into account aspects found in other approaches to political psychology, especially the neo-Freudian (psychodynamic) and the attitude-structural approaches. All three approaches have been utilized to describe and explain the nature and development of the political competencies, and, as we argue below, their contributions to the theory of democratic personality may be integrated into a unifying model. Whether this is a fruitful endeavor should become clear when the empirical hypotheses derived from it are tested.

Psychoanalytic Approach

In the 1930s and 1940s, political psychologists were concerned mainly about the rise of fascism and other totalitarian ideologies. Explanations were given in terms of psychoanalytic theory. The psychoanalytic approach was prominently adhered to by the Berkeley Group (Adorno et al., 1969) which in its study of the authoritarian personality assumed a high degree of coherence between personality and political orientation. For instance, Adorno and his colleagues believed that, if the personality structure lacked “the integration between the moral agencies by which the subject lives,” authoritarian submission, conventionalism, and authoritarian aggression would result and would clear a path for the spread of Fascism (p. 234). Such a “weakness in the ego is expressed in the inability to build up a consistent and enduring set of moral values within the personality, and it is this state of affairs, apparently, that makes it necessary for the individual to seek some organizing agency outside himself” (p. 234). Moreover, there seemed to be a close relationship between authoritarian “adherence to substitutes and crutches of this kind” and immature cognitive organization. The results of their empirical investigations showed that reliance on an “agency outside” relates to “a simple, firm, often
stereotypical structure. There is no place for ambivalence and ambiguities” (1969, p. 480). Though the significance of these findings has been widely acknowledged, the research of the Berkeley Group provoked criticism because of its methodological and theoretical shortcomings. Thus Selznick and Steinberg (1969) argued that a false conception of the role of cognitive factors in political socialization had led Adorno and his associates to an erroneous assumption about the direction of causality in the functioning of personality. Selznick and Steinberg pointed out that fascist attitudes are a result, not a cause, of cognitive structure. Sanford (1973, p. 167), reviewing the critique of research on the authoritarian personality, conceded that the “personality syndromes most useful in understanding political behavior will surely embrace both cognitive and psychodynamic factors.” Such an understanding does indeed seem necessary. More recent psychoanalytic research suggests that neurotic symptoms, which may in part be responsible for the democratically immature personality, are due to a lack of coping ability in the ego rather than, as Freud believed, to an overdeveloped superego (Mowrer, 1972, p. 350). Intolerance of ambiguity, fear of failure, and the feeling of being controlled by external forces may be “type-2-symptoms” (Mowrer), i.e., defensive reactions to an overwhelmingly difficult life situation when the individual lacks the competence to integrate different demands. One might hypothesize that the development of an authoritarian character can be meaningfully described in terms of cognitive-moral development. Kohlberg (1964, p. 422) reported that authoritarianism, as measured by the F-scale of Adorno and his colleagues, correlated negatively ($r = -.52$) with cognitive-moral development. Whereas Kohlberg (1964, p. 422) is reluctant “to offer a view of moral ideology which combines personality type and developmental considerations within a single framework,” other researchers have attempted to do just this (Habermas, 1976; Loewinger, 1976; Döbert and Nunner-Winkler, 1975). For Habermas (1976), Kohlberg's and Loewinger's theories have provided a basis for positive definition of the properties of a democratic personality, an endeavor that Adorno had deliberately eschewed for fear of “false positivism.” One might conclude that, in order to become a democratic personality, one must have a strong ego involving among other things tolerance of ambiguity, hope for success, and an internal control cognition (Lane, 1962, pp. 400-412). Yet this description remains incomplete, if for no other reason than that it could also fit the outline of a narcissistic personality who plays “hardball politics” (Etheredge, 1979). To define democratic personality, we need to take into account both political and moral orientations. [57]
Within the last decade a second major stream of theorizing has merged with the cognitive-developmental approach, resulting in a more unified paradigm of research into political socialization. The object of this research could be labeled “attitude structure” although several other, apparently interchangeable, labels are also in use, e.g., “focus of concern” (Campbell et al., 1960, p. 188), “ideology” (McClosky, 1964, p. 362), “belief system constraints” (Converse, 1964, p. 207), and “levels of conceptualization” (Converse, 1964, p. 215).

The application of the concept of attitude structure to socialization research has corroborated some straightforward hypotheses. As a major result of their study of pupils from grades 2 to 8, Hess and Torney (1967) found that at an early age children acquire some vague “ideal standards” with which they, like adults, evaluate political objects. However, children lack many of the logical, conceptual, and sensory links that make up an integrated self – be they links between these ideal standards and other standards, or between ideal standards and specific beliefs (“issue beliefs”). Many empirical findings show that the process of attitude formation continues beyond childhood and that many adults may never reach the most sophisticated levels of conceptualization, ideology, attitude structure, or belief system constraint (Campbell et al., 1960; Converse, 1964, 1970; McClosky, 1964). Only a small proportion of the population has organized the world of political problems into systems of values and attitudes.

These findings have not been uncriticized. In some studies, belief system constraint is operationalized as “level of verbal conceptualization” (Campbell et al., 1960; Converse, 1964), measuring to a large extent the level of verbal articulateness. This approach may largely underestimate the basic citizen's political competence (Brown, 1970). On methodological grounds, furthermore, there has been objection to the use of (synchronous or diachronous) interindividual correlations as indicators of the degree of individual attitude structure.

On theoretical grounds it has been argued that mainstream attitude structure research has been biased toward the specific “political logic” of political elites, who are found to be more consistent in applying general principles to specific issues (McClosky, 1964, p. 366). This argument neglects the possibility that the man in the street may use different dimensions and ways of organizing attitudes and beliefs. Indeed, Lane (1962) found that the ordinary man also links events and beliefs to form an ideological system, even though this system may be different from the ideology of those who are politically active.

The implicit assumption that the organization of attitudes along a single universal dimension can be regarded as superior to other types of political consciousness now seems doubtful on normative grounds (Lane, 1973; Bennett, 1975). The public may have stable opinions on various matters without
holding to an encompassing ideology. It appears problematic to infer an ill-structured consciousness from a lack of constraint in political attitudes, since a low degree of constraint may indicate a highly structured, highly differentiated consciousness. Even among the political elite, the priority and sufficiency of the liberal-conservative dimension is contested. Furthermore, it is also questionable whether a variety of socio-political value systems can be condensed into a single dimension (Rokeach, 1973; Sandberger, 1979).

To understand political consciousness, one must survey the structure of at least three fundamental value orientations. Ever since the French revolution, liberty, equality, and brotherhood or solidarity have been regarded as core democratic values. Tomkins (1965) asserts that a general humanistic orientation in particular provides the resonance basis for democratic personality structures. He sees the humanistic posture centering on the belief in human goodness, whereas distrust, or sociophobia, constitutes an ideo-affective posture to which anti-egalitarian and authoritarian ideologies tend to resonate. For a heterogeneous sample of respondents, Tomkins (1965) was able to demonstrate a consistent pattern of correlations between expressed empathetic affects and ideological beliefs with regard to a wide range of topics. Furthermore, it seems that a democratically mature value system cannot coexist with political apathy. As Durio (1976) has put it, “democratic behavior is a conscious commitment to a value structure requiring individual action” (p. 212). The “New-Left Ideology” strongly emphasized a more direct and more encompassing participation by ordinary people in political decision making. Thus, a democratic personality is not completely described by value concepts (see Döbert & Nunner-Winkler, in this volume). We have not yet touched upon the more specific behavioral implications of such ideals. Do values exert any influence on specific action decisions? The problem of relating abstract moral and political ideals to specific judgments has been most thoroughly studied by cognitive-developmentalists.

Cognitive-Moral Approach

Cognitive theory of moral development is concerned with ego development, a topic which has long been neglected by psychoanalytic theory and research. According to Kohlberg (1964) one has to interpret “moral character as ego rather than superego strength [. . .] This interpretation implies that the major consistency in moral conduct represents decision making capacities rather than fixed behavior traits” (p. 391). We have seen that this view extends psychodynamic theory. Moreover, the cognitive-developmental approach is concerned with the structural organization of affects which attitude research has usually bypassed for lack of appropriate models and methodological tools.
Kohlberg has described six types of moral judgment which are combined into three levels: preconventional, conventional, and postconventional. On the grounds of moral-philosophical considerations and empirical findings, Kohlberg asserts that these types form an invariant sequence of stages of individual development. He conceives of them as stages of justice and social perspective taking. Moral reasoning is conceived of as hierarchically organized into structural wholes. Invariant sequential order and structural wholeness are important assumptions of the cognitive-developmental approach (Colby et al., 1983; Kohlberg et al., 1983; for a discussion of Kohlberg's theory, see Lind, Chapter 2 in this volume). Theoretically, the degree and kind of structural integration of behavior are the core features of personality development (cf. Loevinger, 1976, pp. 54-67). Hence, we would not expect consistency of judgment to be invariant across situations and throughout individual development.

Methodologically, the assumption of structural wholes implies that we may not infer the level of an individual's cognitive-moral development from a single behavioral event, such as a single answer to a questionnaire item. Multiple assessment of behavior is necessary, not to estimate a "true" attitude score, but to provide information about the structural properties of an individual's behavior. To assess unequivocally the meaning of behavior, we must examine a configuration of responses to a carefully selected pattern of stimuli (cf. Lind 1982b; 1984).

Moreover, the concept of consistency itself carries an ambiguous meaning which must be explicated in every instance. *Behavior is not consistent per se, but always with regard to some value criterion.* Regarding judgment behavior in moral situations, such criteria may be acquiescence (agreement with any argument to avoid debating), opinion agreement (supporting arguments that are in line with a person's issue beliefs), or orientation toward the moral quality of the arguments. Empirical investigations have demonstrated that the degree to which moral principles are consistently applied to judgments and behavior decisions follows some developmental trends, and that this degree also relates to different degrees of political articulation. Conversely, research has indicated that acquiescence and opinion agreement are negatively related to moral development (cf. Fiskin et al., 1973; Keasey, 1974; Lind, 1978a). Only when the subjective validity of moral values becomes independent of egocentric motives, i.e., when the individual is able to *decenter* (Piaget), is the individual's moral competence transformed into what Habermas (1973) calls communicative competence. Only then can individual actions be rationally justified and criticized in terms of principles. The subjective reasoning becomes objective and open to argument from others.
Toward an Integration of Models of Democratic Personality

We have seen that all three of the above-mentioned approaches are structuralist, psychodynamic theory being the least explicit, and cognitive-developmental theory the most explicit in this regard. All three approaches assert that individual behavior and thought are organized into structural wholes, which does not imply that people necessarily behave uniformly across situations. The decisive criterion of belonging to a structural unit is not phenomenal similarity but functional correspondence: Any behavior that serves the same goal can be regarded as an element of the same structure.

Structural wholeness is no metapsychological dogma. It is not a question of “being there or of not being there.” Not every belief or action of an individual is integrated under a single system of orientations, since several reference systems can coexist within the same person without being interrelated. As Adelson and O’Neill (1966) have found, domains of structurally organized behavior in young children may be scattered like islands in the stream of action. Only to the extent that a person has developed consistent relations between various orientations, values, and goals, can he or she be considered to have an integrated personality.

In the process of development, people experience phases of crisis and phases of accelerated change of thought and behavior. From the structural point of view, one is led to believe that this phenomenon is caused by the integration of previously unrelated substructures of personality. In phases of crisis, especially when an individual’s ecological context changes or enlarges, the individual becomes aware of conflicting values and intentions. New criteria have to be found for deciding value priorities, and new modes of behavior have to be acquired to cope with situations of conflict. Contrary to associationist positions, all three approaches take into account both motivational and cognitive components of behavior, although it is not always obvious how these components relate to each other and to behavior. In the face of the widespread practice of treating motivational and intellectual functions separately, one might well ask, “One psychology or two?” (Kuhn, 1978). Though the necessity of combining “a developmental approach with simultaneous interest in motivational and cognitive aspects of personality” (Loevinger, 1976, p. 101) has long been felt, a truly integrated model is still absent (see Kuhn, 1978, p. 116).

In all three approaches, motivational and cognitive components may be conceived of as entities of the same type. Psychodynamic approaches tend to assume that id, ego, and superego are different faculties of the mind which contended with each other for power over behavior. In attitude theory, even approaches that consider the existence of more than one affective dimension conceive of these dimensions merely as juxtaposed to one another. In Lasswell’s (1951) or Greenstein’s (1968) conceptions of the democratic personality,
the cognitive component is essentially restricted to a belief in the basic good-
ness of man, and hence lacks the structural feature under discussion here.

Even in cognitive-developmental theory, the genuine structural character
of moral consciousness is missed when one assumes that the cognitive compo-
nent can be validly measured solely through logical tasks, or that this compo-
nent precedes moral reasoning (Kohlberg, 1976; Kuhn et al., 1977; Ijzen-
doorn, 1979). This conception is questioned by Kaern (1978, p. 98), who
points out that “stages of logical thought and stages of moral thought may be
models of the same theory,” i.e., that the two content areas may have common
structural properties. Such a conception was also proposed by Lee (1971, p.
101), in whose view the ability to see the weight of a body unchanged by the
alteration of its shape seems structurally equivalent to the ability to abstract
the validity of a democratic value from specific applications. Both abilities are
said to be based on the common structural property of what Piaget has called
“conservation.”

Moral behavior always contains aspects of content and of structure which
reflect the distinction between elements and their interrelations. In organismic
structures (see Werner, 1957), elements must be further differentiated into
means and ends, that is, into concrete moral judgment behavior and abstract
value orientations (content). The structure is then characterized by its content,
i.e., the purpose it serves, and by the development of its system of “logical rel-
ations.” As a preliminary operational definition, any value is said to be cogni-
tively organized if it exerts noticeable influence on the pattern of judgment
behavior, i.e., if the value orientation is conserved across a class of judgment
situations. Above all, moral values should be generalized across the particular
population of persons that support them. Each person should be able to take
the value perspective of any other person, which means that the conservation
of values is the core of role-taking ability.

The consistency or “conservation” of moral judgment is a prerequisite for
democratic discourse. An autocratic social system is characterized by the li-
mits it puts on the universal application of values through social power diffe-
rences. In autocratic societies, members of a particular social class are allowed
and even encouraged to prevent the universal application of certain moral
values in order to defend their particularized positions. In such societies, mo-
rality is segmented (for the concept of segmentation, see Senger, in this
volume, and Döbert & Nunner-Winkler, 1975). Conversely, democratic society
rests on the unconditional application of certain basic values. Only if conflicts
do not involve disagreement with regard to basic values can they be solved by
means of rational discourse. Democracy is achieved insofar as individual
judgment behavior is moralized, i.e., insofar as it reflects “all values involved”
(Mead, 1967).
Certain forms of segmentation are, however, legitimate. Considering the availability of means and mitigating circumstances or weighing the values against one another may lead to differentiations in judgment which are compatible with a democratic personality. The opposite monomaniac adherence to abstract values may lead to Jacobinism. A “temporary Jacobinism” is part of most individuals’ adolescence (Lipset, 1965). Perry (1970) conceives of this phase as one of “basic duality,” in which the world is seen in dualistic terms, right versus wrong. But this phase is eventually challenged by relativistic thinking, and finally subdued by the achievement of “contextual relativism,” on which adult “commitment” is based. This commitment brings back the capability of firm judgment, tempered now by full awareness of uncertainty and relativism, “it is an act in an examined, not in an unexamined, life” (Perry, 1970, p. 136). In short, abstract value orientations, like id-impulses and superego-controls, or like general affect-loaded attitudes and goals, are the leading elements of the complex cognitive structure that organizes concrete judgment behavior and issue beliefs. Besides the many things that distinguish the three theoretical “dioceses” discussed above, these contiguities constitute the identification of three defining aspects of personality structure: abstract (motivational) content, concrete (behavioral) elements, and relational (cognitive) structure.

### Table 1

<table>
<thead>
<tr>
<th>Theoretical Domain</th>
<th>Abstract Content</th>
<th>Concrete Elements</th>
<th>Relational Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psychoanalytic Theory</td>
<td>Id, Superego</td>
<td>Thoughts, Dreams</td>
<td>Ego</td>
</tr>
<tr>
<td>Attitude Structure Theory</td>
<td>Affective Component, Orientations</td>
<td>Behavioral Component, Issue Beliefs</td>
<td>Cognitive Component, Constraints</td>
</tr>
</tbody>
</table>

### Research Questions and Hypotheses

This attempt to provide an integrated structural approach to the study of democratic personality and the outline of a model of democratic personality
Democratic Personality

raises several questions, some of which can be investigated empirically. For reasons of time and space, we shall concentrate here on those aspects that seem most fundamental in the sense that they prepare the ground for further research into the nature and development of the democratic personality. *Hypothesis 1.* Prevailing methods of personality and attitude assessment presuppose that people vary only by the amount or degree, not by the structure, of a trait. Inconsistencies in judgment behavior are attributed solely to measurement error. This position of classical test theory (Spearman, Gulliksen) is incompatible with structural theories, which assume that judgment consistency regarding some organizing perspective depends on a person’s cognitive development. We expect, therefore, that even within a developmentally homogeneous group individuals will vary markedly in judgment consistency with regard to particular moral orientations.

*Hypothesis 2.* Judgmental consistency (as defined above), seems to signify moral competence; i.e., we believe that it reflects a particular organization of moral values. This particular organization is primarily defined by the developmental order of moral orientations, as has been suggested by Kohlberg (1971) and Loevinger (1976, pp. 27-28). In operational terms, we expect that the more a person judges concrete statements by the value orientation the statements represent, the more he or she will evaluate them in accordance with a hierarchy of moral orientations as described by Kohlberg’s (1969) stage model.

*Hypothesis 3.* Though the mechanisms are not yet known in detail, we deduce from psychodynamic theory and from cognitive-developmental theory that ego strength and moral competence are correlated. A mature hierarchy of goals and values and their efficient organization, which is reflected in high judgment consistency, should be negatively related to intolerance of ambiguity, fear of failure, and perception of one’s fate as controlled by external forces. Since moral competence means the ability to cope with moral conflicts, it renders superfluous the defense mechanisms of a weak ego, of which those three reactions may be symptomatic.

*Hypothesis 4.* Finally, it is hypothesized that adherence to abstract democratic orientations relates both to the structure and to the content of concrete judgments. Since, as many writers contend, democratic values are among, or even identical with, the highest stages of moral orientation, we expect that explicit commitment to democracy is positively related to moral-cognitive development.

For the present study, we expect that these hypotheses will be empirically valid even when important variables like level of formal education and cohort membership are controlled, i.e., when relationships are tested in a developmentally and socially homogeneous sample. Studies which use more heterogeneous samples will presumably produce even clearer confirmations of these hypotheses.
Methods

Sample

The 708 subjects of this study are a sample token from a cohort of graduates of German upper secondary schools in Baden-Württemberg. Their median age was 18.5. Standardized questionnaires were administered during classroom hours in 48 classes shortly after the students had taken their final (written) examinations in the spring of 1976. This survey is part of a more comprehensive research project on university socialization.

Moral Judgment Competence

For measuring motivational (or evaluative) and cognitive components of moral judgment, a new instrument was developed which utilizes questionnaire techniques and an experimental, multifactorial design. It is called the “Moralisches Urteil Test” (MUT; moral judgment test; Lind, 1978a, 1984, and Lind & Wakenhut in this volume). Its purpose is to infer behavior-guiding orientations and thought organization from an individual's pattern of judgments on the acceptability of certain arguments which are presented in connection with a moral dilemma. The test consists of two subtests, each containing a story which presents a behavioral dilemma, followed by a set of questions pertaining to agreement/disagreement with a suggested solution to the dilemma and to the acceptability of arguments speaking for and against this solution. The arguments were constructed to represent moral reasoning at each of the six stages described by Kohlberg (1969, p. 376). The respondent is asked to evaluate twelve reasons for each story – twenty-four altogether. Hence, by virtue of test construction, the MUT items constitute an “experimental questionnaire,” a tool which has been discussed at length elsewhere (Lind, 1982; see Lind and Wakenhut, this volume). The independent variables form a $2 \times 2 \times 6$ factorial design. The three factors included are: Dilemma, Pro-Con, and Stage. Each item belongs to a particular story, represents a pro or a con position, and refers to one of the six Kohlbergian stages of moral reasoning. The dependent variable is represented by the respondents' judgment of the acceptability of the given reasons on a scale ranging from -4 to +4. Since neither self-description nor introspection is recorded, only individual patterns of judgment behavior, this questionnaire can be viewed as a behavioral test. The dilemma of the subtest entitled “Mercy Killing” is adapted from Kohlberg's Situation IV (Kohlberg, 1958, p. 366). It is assumed that the life-death issue requires the most unambiguously moral reasoning. On contrast, we used a second dilemma labeled “Theft,” a story
Democratic Personality

about two workers who break into the main office of their factory in search of evidence to support the allegation that the management has bugged their work area (for more details on the MUT, see Lind & Wakenhut, in this volume).

In this paper the inter-stage structure is elicited by calculating the average acceptability of every four statements for each of the six stages, and by combining them into individual profiles of stage-ratings. This method is a convenient device for depicting simultaneously the motivational (content) and the cognitive (structural) components of moral reasoning (cf. Lind, 1985a). The content of the judgment is indicated by the acceptance or rejection of the stage of orientation which the argument represents. The degree of structural or cognitive organization of judgment behavior is indicated by two measures: the distance of the average acceptability of a stage from the theoretical scale mean (±0) and the steepness of the profile of stage-ratings. The less consistently a respondent evaluates statements with regard to their stage-appropriateness, the more the profile flattens and approaches the zero-line. In addition to these profiles, a direct measure of intra-individual response consistency is used, although this measure by itself does not allow us to infer content from structure.

Ego Strength

Three scales for measuring ego strength were included in our study. Intolerance of ambiguity was measured through the scale developed by Budner (1962), which seemed appropriate since it is explicitly related to psychoanalytic concepts. The scale contains 16 items which are to be rated from -3 (“I reject completely”) to +3 (“I accept completely”). After a linear transformation, the total scale ranges from 0 to 96 with 96 signifying extreme intolerance of ambiguity. Fear of failure, which is considered either to result from ego-weakness or to contribute to it, was assessed using seven items adapted from a larger Fear of Failure scale developed by Fend et al. (1974). The respondent was to decide between pairs of statements, e.g., “When I am confronted with a new task … (A:) I am rather sure that I shall succeed. (B:) I am often afraid that I shall not succeed.” The scale ranges from 1 to 7, with 7 indicating extreme fear of failure. Locus of Control was measured by a shortened version of Rotter's (1966) scale. Our selection of items represents the personal control factor (cf., Gurin et al., 1969; Lefcourt, 1976).
Democratic Orientation

For this study we selected four indicators of democratic orientation. The questions we chose for the present analysis pertain to the following issues. Respondents were asked to take a stand on two questions involving egalitarian values. First, “Would you be for or against a reduction of social inequalities?” and second, “Would you support an increase of social equality even at the cost of material wealth?” These items are part of a comprehensive instrument for the study of orientations toward social inequality (Sandberger, 1983). For assessing Humanism a somewhat shortened version of Tomkins’ (1965) scale was used. The respondent was to choose between polar statements, e.g., “Human beings are basically good” versus “Human beings are basically bad,” with the option to indicate neutrality or indifference.

Democratization, taken as an abstract ideal of political action, is described by an item of the “New Left-Scale” developed by Christie et al.: “The democratization of all areas of life should be the basis of a new society” (cf. Gold, et al., 1976; Robinson & Shaver, 1973, p. 470). The respondent was to indicate his degree of agreement or disagreement on a scale ranging from -3 to +3. A readiness to participate in political affairs was seen as a prerequisite for democratic orientation. We took as an indicator of this orientation the respondents' reaction to the statement, “I will not engage in political activity under any circumstances.”

Evaluation of Data

To gauge the degree to which the data support our hypothesis, we have used two complementary approaches. In the first approach, information content and conformity with theoretical prediction are tested, and in the second statistical significance is checked. Our hypothesis is that students high in ego strength and democratic orientation will discriminate more between stages of moral reasoning, rejecting lower-stage arguments and accepting higher stage arguments more markedly than do their counterparts who are not so high in those orientations. In technical terms this means that the former group should exhibit a steeper profile of median scores for the stages. A thought experiment reveals that, for this hypothesis, the a priori chance of confirmation is rather low. When the median scores of two groups – to be denoted here as A and B – are compared for each stage, and a dichotomous distinction is made between “A smaller than or equal to B” and “A greater than B,” there are \(2^6 = 64\) possible sequences or configurations across the six stages. Among these, six configurations are in line with the hypothesis; their graphic form corresponds to that of an X. Hence, under the assumption of randomness and independence of trials (stages), the probability of obtaining a configuration that conforms to
the hypothesis is no larger than 6/64 or about 9 percent. The reciprocal value of this number constitutes a measure of the information contained in the data (see Lind, 1984a).

For computing statistical significance, two-way analyses of variance were run, using ego strength and democratic orientation – taken one at a time – as independent variables, and using stage as a repeated measurement variable. To analyze the effects of each of these variables, orthogonal polynomials as proposed by Bock (1975, pp. 447-488) were used. Our hypotheses focus interest on the interactions between the variables of ego strength and democratic orientation with stage, with special emphasis on the linear component of this interaction. To support the hypothesis of a differential steepness of stage-score profiles, the linear component should be significant and should account for a sizeable proportion of the variance.

Empirical Findings

Hypothesis 1: Variation in the Cognitive Aspect of Moral Judgment

Intra-individual analysis of components of variance calculated from the MUT reveals that individuals do indeed differ greatly with regard to response consistency or, more precisely, consistency with respect to moral criteria. When confronted with a behavioral dilemma, some persons employ moral reasoning to a high degree while others consider nonmoral criteria of judgment, e.g., whether the reasons advocated support, or oppose, the opinion to which the person is committed. The degree to which respondents use moral categories in their judgments is depicted in Figure 1. The abcissa represents the proportion of individual judgment variance accounted for by the Stage factor. This measure is simply the ratio of the sum of squares due to this factor to the total sum of squares. In Figure 1, the degree to which individual judgment behavior is determined by the opinion-agreement variable (Pro-Con) is also depicted as well. As compared to younger subjects (cf. Keasey, 1974), the upper secondary school graduates in our sample are obviously less oriented toward the agreement/disagreement aspect of the arguments.
Hypothesis 2: Cognitive and Affective Aspects of Moral Judgment

If the preference for "high" levels of moral orientation is a pacer for a later form of cognitive organization, we are led to assume that the recognition of the moral priority of higher stage reasoning should go together with a more highly organized thought structure, i.e., that motivational content and cognitive structure of moral judgment behavior should empirically correlate in a predictable way. The findings depicted in Figure 2 clearly corroborate this hypothesis. The more the "higher" moral orientations are accepted, and the more the "lower" stage reasons are rejected, the more the respondents evaluate the arguments morally, i.e., with regard to the stages they represent. Configurations of intergroup median differences are all perfectly in line with the prediction. Thus this finding supports the claim that even Stages 5 and 6 have a cognitive basis, a claim which has sometimes been questioned (cf. Gibbs, 1977, among others).

Hypothesis 3: Moral Competence and Ego Strength

A marked and well-structured value hierarchy coincides with high ego strength. Though the sample as a whole can be regarded as comparatively mature and as fairly homogeneous in its level of moral development, differences in ego strength correlate with differences in moral judgment structure. Although the orderings of moral orientations (stages) are similar, students of lo-
Democratic Personality

Wer ego strength are obviously less consistent in the application of abstract orientations to concrete judgments. This is indicated by a less marked rejection of lower stage arguments. The pertinent data are presented in Table 2. High scores on intolerance of ambiguity turn out to be associated with relatively high endorsement of Stage 1 through Stage 3 reasoning. With regard to fear of failure and external control cognition, the findings are somewhat less articulate, yet they follow the same pattern. Those who fear failure, like those who feel they are externally controlled, exhibit a less consistent rejection of morally inferior orientations in their judgments. Except for the comparison of students who are medium and low in intolerance of ambiguity, all configurations of median differences between groups are fully congruent with the hypothesis. In each case the dependent analysis of variance yields a significant linear component of the Groups x Stages interaction; although, for intolerance of ambiguity, some higher order components also turn out to be significant, the bulk of the interaction variance is accounted for by the linear component, as shown in the last column of Table 2. This confirms the hypothesized differences in steepness of stage rating profiles.

These findings corroborate those of other studies. Sullivan and Quarter (1972, p. 156) report that Conventionals reveal a lower, “morally hybrid” and “pure” Postconventionals a higher, tolerance of ambiguity as measured by the Omnibus Personality Inventory. Haan et al. (1973) found a low but consistent correlation between tolerance of ambiguity and Kohlberg's moral maturity measure. Alker and Poppen's (1973) study showed that moral development was negatively related to dogmatism. Correlations of moral development with the locus of control variable are only partially comparable, since in most cases the complete Rotter scale was used, which confounds several dimensions (Mirels, 1970, Gootnik, 1974). Nevertheless, (low) negative correlations between moral development and external control cognition have consistently been found (Alker and Poppen, 1973; Bloomberg, 1974).
Finally, we hypothesized that conscious adherence to democratic values would correlate with moral-cognitive structure. The results shown in Table 3 are in line with this hypothesis. Though egalitarianism is only a prerequisite concept and not necessarily indicative of a fully developed democratic consciousness (Piaget, 1973; Rawls, 1971), students who support a reduction of social inequality – even at the cost of material wealth – show less preference for lower stage reasoning. This can be taken as an indicator of higher moral development. Humanistic values and the request for “democratization of all areas of life” seem to be at the heart of advanced moral development. This is indicated by the preference profiles of the “high” and “low” groups as presented in Table 3, and by their differences with respect to cognitive-moral consistency. In our sample, students with humanistic and democratic orientations show the clearest preference order of the six stages of moral orientation. The group of students disagreeing with democratization shows less discrimination between the stages of morality in evaluating arguments. In other words, this latter group possesses less decisional competence for making morally substantiated judgments. However, this is a small group. Except for egalitarian attitudes, the configurations of median differences conform perfectly to the hypotheses. Analyses of variance yield significant linear components for the Groups x Stages interactions for all the aspects of democratic orientation examined here. With respect to some of these variables, higher order components of the interaction also turn out to be significant, yet in all cases but one the linear component accounts for the greater part of the interaction variance.
Conclusion

A central prerequisite for a democratic society is the moral autonomy of its citizens. A person may be regarded as morally autonomous to the extent that he or she exhibits “general consistency of approach on principle to all situations” (Kohlberg, 1958, p. 131). This does not mean, of course, cross-situational rigidity. It means a greater awareness of the multiplicity of aspects of a situation, and a more general and more encompassing perspective on moral values. All three structural approaches discussed in this paper agree that a complex cognitive structure is needed to organize behavior in accord with such abstract ideals. As Nunner-Winkler (1980) has asserted, the stability and maintenance of democracy is intimately linked to the degree to which moral values are structurally anchored in the individual mind. It obviously does not suffice to reach consensus on the basic democratic values (Dahl, 1961, p. 325). Inability, or unwillingness, to apply these values in concrete situations renders the individual prey to autocratic submission.

Our integrated study of styles of reasoning, ego functioning, and values suggests that ego strength (tolerance of ambiguity, hope for success, internal control cognition), content and structure of moral judgment behavior, and positive valuing of democratic ideals are all interrelated. These findings, as well as those from other studies, provide insight into the importance of the cognitive component in democratic behavior. Hence, they clearly demonstrate that, as Binford (1983) has stated, “the tolerance of opposing viewpoints and the appreciation of the utility of such views for social change is an important organizing principle for the character-rooted democratic personality” (p. 678). Moreover, they suggest that the acquisition of abstract values may be considered a necessary, though not sufficient, condition for the development of democratic personality. Further studies are needed to clarify how value ideals function as pacers for cognitive development.
### Table 2

**Moral Judgment and Ego Strength: Median Stage Acceptability by Students with Different Scores on Ego Strength Scales**

<table>
<thead>
<tr>
<th>Group</th>
<th>( n^b )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>( (c) )</th>
<th>( (d) )</th>
<th>( (e) )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intolerance of Ambiguity:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low</td>
<td>(0–27)</td>
<td>272</td>
<td>-2.1</td>
<td>-1.8</td>
<td>-1.7</td>
<td>-0.4</td>
<td>0.9</td>
<td>0.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>medium</td>
<td>(23–43)</td>
<td>411</td>
<td>-1.4</td>
<td>-1.4</td>
<td>-0.9</td>
<td>-0.1</td>
<td>0.7</td>
<td>0.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>high</td>
<td>(46–96)</td>
<td>25</td>
<td>-1.1</td>
<td>-1.4</td>
<td>-0.3</td>
<td>0.3</td>
<td>0.6</td>
<td>0.6</td>
<td></td>
<td>1% 75%</td>
</tr>
<tr>
<td><strong>External Locus of Control:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>internal</td>
<td>(6–11)</td>
<td>200</td>
<td>-1.9</td>
<td>-1.8</td>
<td>-1.4</td>
<td>-0.3</td>
<td>0.8</td>
<td>0.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>external</td>
<td>(12–18)</td>
<td>431</td>
<td>-1.6</td>
<td>-1.4</td>
<td>-1.1</td>
<td>-0.2</td>
<td>0.8</td>
<td>0.8</td>
<td></td>
<td>1% 94%</td>
</tr>
<tr>
<td><strong>Fear of Failure:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low</td>
<td>(1–3)</td>
<td>250</td>
<td>-1.9</td>
<td>-1.8</td>
<td>-1.4</td>
<td>-0.2</td>
<td>0.8</td>
<td>0.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>high</td>
<td>(5–7)</td>
<td>324</td>
<td>-1.5</td>
<td>-1.4</td>
<td>-1.2</td>
<td>0.2</td>
<td>0.8</td>
<td>0.8</td>
<td></td>
<td>1% 79%</td>
</tr>
</tbody>
</table>

\( a \) From -4 (completely unacceptable) to +4 (completely acceptable).
\( b \) The \( n \)'s usually do not add up to the sample size of 708 because of "don't know" answers.
\( c \) Configurations of median differences that are completely in line with the hypothesis are denoted by brackets.
\( d \) Statistical significance of the linear component of the Group x Stage interaction. Two-way analysis of variance with Stage as a repeated measurement variable, applying orthogonal polynomials according to Bock (1975, pp. 447–488).
\( e \) Proportion of the total interaction variance accounted for by the linear component.
Table 3

Moral Judgment and Democratic Orientations: Median Stage Acceptability by Students with Different Scores on Dimensions of Political Attitude

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>Acceptability of Stage</th>
<th>Testing of Hypotheses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td><strong>Reduction of Social Inequality:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>For</td>
<td>563</td>
<td>−1.8</td>
<td>−1.6</td>
</tr>
<tr>
<td>Against</td>
<td>92</td>
<td>−1.4</td>
<td>−1.3</td>
</tr>
<tr>
<td><strong>Increase of Equality:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Support</td>
<td>391</td>
<td>−2.0</td>
<td>−1.6</td>
</tr>
<tr>
<td>Reject</td>
<td>126</td>
<td>−1.4</td>
<td>−1.3</td>
</tr>
<tr>
<td><strong>Humanism:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>620</td>
<td>−1.7</td>
<td>−1.5</td>
</tr>
<tr>
<td>Low</td>
<td>51</td>
<td>−1.1</td>
<td>−1.2</td>
</tr>
<tr>
<td><strong>Democratization:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>437</td>
<td>−1.9</td>
<td>−1.6</td>
</tr>
<tr>
<td>Disagree</td>
<td>46</td>
<td>−1.4</td>
<td>−1.1</td>
</tr>
<tr>
<td><strong>Participation in Politics:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>409</td>
<td>−1.9</td>
<td>−1.7</td>
</tr>
<tr>
<td>Not as all</td>
<td>72</td>
<td>−1.2</td>
<td>−1.1</td>
</tr>
</tbody>
</table>

Note: For explanation see the notes of Table 2.
With the combination of survey research methods and sophisticated assessment techniques, we hope that new perspectives will open up for the study of processes of socialization and personality development. The new techniques for assessing moral judgment structure render cognitive-developmental theory, as well as other structural theories, testable through empirical research geared to using multiple measures with large samples of individuals. Moreover, it will be possible to systematically investigate processes of moral décalage, value conservation, cognitive anchoring, or conversely, processes of moral segmentation, particularization, and isolation. This research will ultimately have to be directed at answering Allport's (1929) fundamental query of whether a fully developed democratic personality is merely a particular configuration of unrelated drives, attitudes, or traits, which are accidentally juxtaposed by an individual's history of conditioning and shaping, or whether such a personality represents a more or less organized whole of thought and behavior, which is continually striving for integrated ego identity.
Democratic Personality

Notes

1. These are not identical with “behaviorist” psychology. Our cognitive-developmental approach is behavioristic in so far as it claims to be empirically. In contrast to the concepts of “latent traits” and “hypothetical constructs,” the organization of cognitive structure requires that it be empirically verified by the structural assessment of patterns of judgment behavior. But cognitive-developmental theorizing is to be distinguished from atomistic behaviorism, which rules out a priori the possibility of measuring complex constructs by means of behavioral assessment. The position here may best be designated by Miller, Galanter, and Pribram’s (1960) term “subjective behaviorism.”

2. See Edelstein et al. (1980), who found that, contrary to Kohlberg’s (1976) assertion, the hypothesis of a “general genetic primacy of logical, before social-cognitive, operations is not tenable” (p. 12).

3. By referring to “logical relations,” we allude not only to propositional logic, on which Piaget has concentrated in his work, but also to the broad, yet not fully recognized, fields of modal and deontic logic. The insufficiency of Piaget’s approach hal in the light of Hegel’s dialectical logic, been thoroughly discussed by Broughton (1981).

4. This study was carried out by the Forschungsgruppe Hochschulsozialisation at the University of Konstanz, as part of an ongoing international longitudinal study. Besides the authors, Barbara Dippelhofer-Stiem, Gerhild Framhein, Hansgert Peisert (director), and Hans-Gerhard Walter have collaborated on the research. The cross-national FORM project is coordinated by the European Coordination Centre, Vienna, and comprises parallel research projects in Austria, the Federal Republic of Germany, the Netherlands, Poland, and Yugoslavia. The research design and questionnaire used with the graduates of West German upper secondary schools (Abiturienten) are fully described in two working documents, prepared by the Forschungsgruppe Hochschulsozialisation: Wissenschaftlicher Bericht, 1976-1978 (Zentrum I Bildungsforschung, University of Konstanz, 1979); and Anlage und Instrumente des Abiturienten-Follow-up, 1976177 (Arbeitsunterlage 37, Zentrum 1 Bildungsforschung, University of Konstanz, 1981).

5. Note that these experiments are located on the level of the individual (idiographic experiment), and not on the level of an aggregate of individuals, as is the case in the correlational analysis below. Nor are these experiments conducted to assess the effect of an “independent” variable, but to measure individual structural dispositions (Lind, 1985d).

6. In formal terms, these configurations consisted of i times A smaller than or equal to B, followed by (6 - i) times A greater than B, where A denotes the median scores of the group higher in ego strength or democratic orientation, and the range of the index i is between 1 and 6. In these configurations, scores A are at most equal to scores B for lower stages, and are larger than B for higher stages. By way of limiting cases, configurations with lower A scores for lower stages and equality of scores for higher stages are included in the conform subset, whereas, of course, the case in which A = B for all stages is excluded.

7. The Box-Bartlett Test (Morrison, 1976, pp. 252-253) serves to check
whether the assumption of homoscedasticity (equality of variance-covariance matrices for independent groups) is tenable. As revealed by the Huynh-Feldt pattern test (Morrison, 1976, pp. 214-215), the model that Bock (1975, pp. 459-460) presents as Assumption 1 proved too restrictive. The tests we applied (using term-by-term F statistics) are valid under Bock’s most general Assumption III, which allows for heteroscedasticity as well as correlated polynomial error components.

8. Analysis of variance procedures cannot be meaningfully applied to these data, because the way the consistency measure was constructed implies that the homoscedasticity assumption is violated, that is, that the groups have different interstage covariance matrices. These are not identical with “behaviorist” psychology. Our cognitive, developmental approach is behavioristic in so far as it claims to be empirical. In contrast to the concepts of “latent traits” and “hypothetical constructs,” the organization of cognitive structure requires that it be empirically verified by the structural assessment of patterns of judgment behavior. But cognitive-developmental theorizing is to be distinguished from atomistic behaviorism, which rules out a priori the possibility of measuring complex constructs by means of behavioral assessment. The position here may best be designated by Miller, Galanter, and Pribram’s (1960) term “subjective behaviorism.”

A preliminary version of this chapter appeared in Political Psychology (Lind, Sandberger & Bargel, 1981-82). We wish to thank the publisher of the journal for her kind permission to revise that article for the present volume.
3 Testing for Moral Judgment Competence

Georg Lind and Roland Wakenhut

Few if any psychological instruments have been based upon such an intimate dialogue between theory formation and empirical research as has Kohlberg's well-known moral judgment interview, especially in its revised, "issue scoring" version. The method embodied in this instrument has proven its worth in numerous studies since the first time it was used (see Kohlberg, 1958; 1984; Colby et al., 1983), and by now belongs among the list of methods firmly established in psychological research. This method, however, is not without controversy, which is hardly surprising in view of the conceptual complexity of many of its central terms, such as “morals” and “competence in moral judgment.” It was therefore only to be expected that alternative methods of investigation would arise, and with them instruments which would attempt to extend or even replace Kohlberg's moral judgment interview.

In this paper we shall discuss some of those new methods that assess competence in moral judgment. What information can be gained with them? How does one proceed in designing questions and answers and statistical methods which are theoretically meaningful? And also, can – as seems to be generally assumed – the classical criteria of test construction be applied to the evaluation of structural personality tests?

Cognitive-developmental theory does not always permit a conclusive decision in favor of one specific operational form in moral judgment research. To justify the development of a particular measuring instrument, further specification is necessary based on the specific scientific issue at stake. Thus the process of constructing objective tests for psychological research (Loevinger, 1957) itself becomes an important means of increasing knowledge. Moreover, we believe there are some methodological postulates inherent in cognitive-developmental theory which call also for new decisions in psychometrics.

Like Kohlberg we want to concentrate on concrete moral judgment behavior. In order to distinguish clearly between this approach and that of questionnaires in which persons have to report about how they “usually” behave or how they would “possibly” behave, we have designated the methods to assess judgment behavior as “Experimental Questionnaires” (Lind, 1982b) or more often, simply as “Tests.” Following the cognitive developmental approach, the response behavior of the persons questioned is not considered to

---

2 For new findings concerning the theoretical and cross-cultural validity of the renamed "Moral Competence Test" (MCT) see:G. Lind (2016). How to Teach Morality. Berlin: Logos, chapt. 4.
be a sign of a merely hypothetical moral disposition; rather, the pattern of judgment behavior is regarded as a manifestation of the moral attitudes and cognitions, which are considered as the two basic defining aspects of moral behavior (cf. Kohlberg, 1979a; cf. ch. 2 in this volume). In this way a number of problems involved with introspective, self-reporting questionnaires avoided – problems that result from the interference of the characteristic under study and the subject's state of being when answering the questionnaire (response set, etc.; see Eckensberger et al., 1980; Lempert, 1982).

Interviews and Tests

The instruments which we shall discuss under the rubric “tests” do not include those methods that merely attempt to approximate Kohlberg's interview method but thereby often ignore the central postulates of cognitive developmental theory. As Schuhler (1978) noted, those methods are inherently unable to account for the essential assumptions of assessing (competence) in moral judgment. Our point of departure is the still stronger claim that there are central statements in cognitive theory of moral development which contradict the “hidden anthropological requirements” of the theoretical assumptions of classical tests and questionnaires.1 Whereas the latter assume that the characteristics of personality can be separated like substances into isolated components, the cognitive developmental theory postulates that the effective and cognitive aspects of moral judgment competence are aspects of the same mental “substance” and thus can only be dealt with together.2

The best way of determining the differences between the Kohlbergian interview method and the tests of moral judgment competence described here is to examine the operations involved in interviewing and evaluating. Both methods have certain features in common. They involve systematically varied testing of moral judgment behavior and are standardized to a great extent in the way they are conducted and evaluated. In both, several dilemmas are presented as short stories and both require of the respondent the ability of judging the adequacy of moral reasons. The major difference is that interviews are production tasks, whereas the tests ask for recognition. In the Kohlbergian interview a number of questions follow each other which the subject has to answer and account for in his own words. (Did Person X act in the right way? Why do you think so? etc.) In the questionnaire, answers and reasons are given to the subject, who must then rank them on the basis of their acceptability. In the interview technique the answers given by the subject are compared afterwards step by step with the exemplars or criterion answers provided in the coding manual. Thus for every issue, e.g., Life, a score is determined from which a total score is calculated arithmetically. In contrast, in tests the “criterion answers”
are presented directly to the subject, who then compares them with his or her own moral judgment. Regarded from a practical point of view, the two methods are not antagonistic; their differences reflect a difference of research interest. Production and recognition are both important expressions of moral judgments, which complement rather than substitute for each other. However when they are regarded from a theoretical point of view, the differences between experimental questionnaires and the Kohlbergian interview are more subtle.

These differences can be set forth in terms of the following four problems associated with the interview method. First, the statistical evaluation of the moral interviews is still virtually atomistic despite great efforts to make it “more structural.” The assignment of a response to a stage according to the coding manual still presupposes a classical theory of psychometrics in which each individual response is taken as a sign of a probable underlying trait. When variation of reasoning across dilemmas and issues occurs, the variation is not, as one could expect on the basis of the structural theory, taken as an important manifestation of organization of judgment but is viewed merely as an error in the assessment instrument (cf. Colby et al., 1978; 1983). Therefore, the calculation of the mean moral judgment scores neglects the structural or process-related information in manifest judgment behavior.

The second problem involved in the interview method has to do with differential motivation in judgment behavior. Interview scores reflect not only a person’s ability to argue at particular stages but also his or her motivation to exhibit “high-sounding” moral arguments in a relatively artificial situation like an interview, in which there is no real or perceived need to reason at the highest stage. This can result in underestimating the stage of moral judgment competence.

Third, in the interview method the force or intensity of an individual’s argument is not recorded. It is only inferred from the number of times it has been given in the course of the interview. The assumption of a close correlation between intensity and frequency is valid in many cases, but it is not always close enough to enable a reliable inference to be made. In contrast, experimentally designed tests, which clearly differentiate between the intensity and the consistency of responses, can provide information on the force of an argument directly, without relying on unsecured inferences. Fourth, if one is interested more in the moral arguments people perceive and prefer, and less in which moral arguments they produce themselves, the interview method would stand second to questionnaires with regard to theoretical validity (Rest, 1975b, p. 748).

In sum, tests – at least tests in the strict sense used here of an experimental
questionnaire (Lind, 1982b) – have some advantages over interviews. They are more economical, a fact which permits their extensive use and the examination of hypotheses requiring great numbers of people; they are more transparent, a fact which promotes a critical examination of their theoretical (content) validity as well as efforts to analyze moral judgment behavior and its determinants differentially; and in tests the intensity of preference for a particular moral orientation can be directly obtained by appropriate response scales. However, this does not mean that tests or questionnaires can completely replace the clinical interview method. Highly standardized and carefully designed interviews are, in many cases, indispensable for research as well as for praxis. The method to use depends, above all, on the particular question one has in mind.

**Tests for Moral Judgment Competence**

Because we did not want to construct only an economical replacement for the interview method, but also a research instrument capable of answering particular theoretical questions in current moral psychology, we have created new “Tests for Moral Judgment” (TMJ), where by “moral judgment” is meant moral (competence) in the same sense described above (n.1). Other questionnaire methods like the Defining Issues Test (DIT) by Rest (1979a) come close to our aims, but they fail to meet the special requirements which we believe characterize cognitive-developmental theory.

Regarding tests of moral judgment, there are two different kinds of procedures which vary according to the method's design and the strategy involved in evaluation: (a) classical tests of moral attitudes and (b) new structural tests for both affective and cognitive aspects of moral judgment behavior. The DIT is the best example of the first type we have in mind. It has proven its worth as a psychometric “technique” in Coombs's sense of the word (Coombs et al., 1975, p. 45); it is problematic, however, as a “criterion” for cognitive developmental research. The DIT is based on Kohlberg’s theory of moral development. At the same time, it complies to the rules of classical testing theory, rules whose basic assumptions seem basically incompatible with the cognitive theory of development. The variation of responses across more than one stage or level (i.e., the degree of inconsistency), is, on the one hand, taken as a methodological concept to determine the state of development, and, on the other hand, used as empirical evidence against the theory. But these two interpretations are mutually exclusive. Rest has resolved this dissonance between theory and method partly at the expense of the theory of cognitive development. Of the two components of moral judgment competence, the DIT deals explicitly
with the content or affective aspect. However, the structural-cognitive component is expressed implicitly by the P score, which is a measure of response consistency in relation to postconventional morality.3

The second type of method pursues and seeks to extend this line of structural diagnostics. It refers to Rest’s earlier research on the relationship between arguments produced on one’s own and preferences for given arguments (Rest et al., 1969; Rest, 1973). In contrast to these pure preference tests, structural tests include the cognitive aspect of moral judgment behavior, in addition to the content-affective components of moral attitudes. Dealing with both components simultaneously seems indispensable, “because a moral act or attitude can be defined neither by purely ‘cognitive’ nor purely by ‘motivational’ criteria” (Kohlberg, 1958, p. 16).

In two instruments, the attempt has been made to realize this new approach. The German (Moralisches Urteil Test) (MUT; moral judgment test), which requires a judgment on acceptability, was developed as part of the research project “University Socialization”4 which accounts for some of its particular features such as the highly differentiated reasons scale, and the selection of themes. The test has been used in studies of various groups of people in Germany, Austria, and Switzerland, and translated versions have also been employed in moral judgment research in the Netherlands, Poland, and Yugoslavia (Lind et al., 1976; Lind, 1978a; 1984b).

The second test, the Moralisches Urteil Fragebogen (MUF; moral judgment questionnaire) developed by Krämer-Badoni and Wakenhut (1978b; Wakenhut, 1982) constitutes a variation of the MUT and has been geared to the special goals of the study set by the authors; it was originally created for the research project “Socialization in the German Armed Forces,”4 but in the meantime has also been used in a number of other studies, as has the MUF. New dilemmas, a modified response scheme, and special characteristic statistics were developed for the MUF – the least of these being especially important in view of the project’s aim to measure the phenomenon of segmentation in moral consciousness (cf. Krämer-Badoni & Wakenhut, and Senger, in this volume).5

The similarities and differences between these three types of methods for measuring moral judgment competence, i.e., the Kohlberg interview method, the Defining Issues Test and the tests for measuring moral judgment competence (MUT and MUF) discussed here, are shown in Table 1.
Design and Construction

The construction of tests to measure competence in moral judgment presents four problems: systematics of test designs, selection and construction of moral dilemma, selection and construction of moral arguments and answers, and computation and interpretation of various indexes from the individual's pattern of judgment behavior.

Experimental design

Since we view tests of moral judgment competence as behavioral tests, which allow us to infer affective and cognitive dispositions of judgment behavior, the question of test design is of particular importance. A person's ability to bring to bear in concrete acts of judgment moral attitudes, which compete with other orientations that can potentially guide action, and to do so in an integrated and differentiated manner, cannot be measured by individual items (or their summation). It must be dealt with in terms of the structural relationship between the pattern of questions and the pattern of answers. In structural methods, systematic questioning facilitates the hermeneutic intention to create the possibility for testing competitive hypotheses about an individual's motive in judgment behavior (see Lind, 1982b; 1984a). This approach can be found in the intuitive assessments of other persons in daily life, and is made explicit by the scientific techniques of standardized interviews and experimental tests. The structure of a TMJ consists of elements (items) and their relations (factorial design). Whereas the items elicit individual behavioral acts, the factorial design is devised to reveal the hypothesized dispositions – when they really exist in an individual. Because a TMJ is designed to measure judgment behavior oriented to moral categories, the items given must represent moral concerns. The MUT and the MUF contain Kohlberg's (1969) six different moral stages as their main design factor. Each type or stage is represented by argumentation which the subject has to evaluate. With this experimental design, one can determine whether moral dispositions are effective in an individual pattern of
<table>
<thead>
<tr>
<th>Method</th>
<th>Behavior</th>
<th>Dispositional Aspects</th>
<th>Cognitive Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kohlberg Interview</td>
<td>Production of moral arguments</td>
<td>Stage-typical orientations, concerns, norms, elements</td>
<td>Consistency of moral argumentation on systematic probing</td>
</tr>
</tbody>
</table>

Measure: The Moral Maturity Score (MMS), common to both aspects and based on frequency of moral arguments produced.

<table>
<thead>
<tr>
<th>Defining Issues Test (DIT)</th>
<th>Judgment of the importance of given moral arguments</th>
<th>Intensity of preference for stage-typical concerns</th>
<th>Consistency of judgments regarding postconventional reasoning stages</th>
</tr>
</thead>
</table>

Measure: The proportion of postconventional preferences in the total of top-ranked arguments, P score, common for both aspects.

<table>
<thead>
<tr>
<th>Test for measuring Moral Judgment Competency</th>
<th>Judgment of the acceptability or legitimacy of given moral arguments</th>
<th>Direction and intensity of preferences for stage-typical concerns</th>
<th>Consistency of judgment behavior regarding one's concerns independent of their opinion-conformity</th>
</tr>
</thead>
</table>

Measure: Different scores for the affective and the cognitive aspects of moral judgment, as well as common scores.
Testing Moral Judgment Competence

judgment behavior in that one assesses the degree to which the person consistently rates arguments according their moral differences.

The TMJs attempt to analyze the degree to which moral judgment is integrated and differentiated – a task for which a one-factorial construction and evaluation design is not sufficient. Whereas in classically designed questionnaires, one is solely concerned with the degree of attitude consistency (integration), the use of multifactorially designed Experimental Questionnaires permits us to investigate the degree of differentiation of a person's judgment. That is, TMJs are intended to assess properties of judgment behavior which in the classical questionnaire construction are covered by the concept of “measuring error.” In latter, error is equated with the total of unknown factors and thus has to a great extent removed these factors from the researcher's attention. Some of these other factors, however, we already know of from cognitive developmental theory, though they have not been the subject of much if any empirical study. In the course of discussing a decision dilemma, agreement or disagreement with a line of argument may be supported not only by one's moral concerns but also – or even exclusively by other motives or combinations of motives. It can result from mere respect on the subject's part for the psychological instrument or from a general tendency to agree with whatever arguments are given (acquiescence); or agreement may depend on whether the argument for one's own standpoint speaks for a solution to the dilemma (opinion conformity).

Acquiescence and opinion conformity are the two most important orientations which have to be considered in addition to moral motive when studying a person's moral judgment competence in order to avoid false interpretations (Lind, 1978a, p. 181-185). If we would study only a person's evaluations of a particular stage of conforming moral reasoning, for example, of “law and order” reasons that support his or her opinion on a moral dilemma, we would not be able to unambiguously infer from these evaluations which of the three motives have determined them. In the TMJs, the moral motive is differentiated from acquiescence by examining the person's evaluations of qualitatively different arguments. (According to the terminology of analysis of variance, ANOVA, we call this the “Stage factor.”) In order to differentiate moral concerns from a person's concern for supporting his or her intuitive opinion, the TMJs incorporate arguments based on both conforming and contrary opinions (the “Pro-Con factor”). Kohlberg also uses this technique in his interviews to test the “firmness of a person's 'structure'” (Eckensberger et al., 1980, p. 340; Kohlberg, 1958, pp. 12 and 131).

As a third dispositional factor to be considered when assessing moral judgment is the person's differential attention to the particular moral principles involved in the moral dilemma that is presented. As Kohlberg wrote in 1958, a morally adequate solution to a dilemma must “do justice to one's own
conviction and the demands of the situation” (p. 128).

To deal with situational differentiation and hierarchical integration of moral judgment in the total personality systems, the TMJs take the dilemma situation into account as an experimental factor (the “Dilemma factor or Story factor”). By combining these three orientation factors, extensive models on the structure of judgment behavior can be tested empirically. We are thereby able to deal not only with the main effects but also with instances of personality integration that go beyond consistency in regard to single moral principles.

As behavioral tests, TMJs do not stop with the assumption of a hypothetical construct, but aim at moral dispositions which are actual – that is, dispositions which are manifested in the act of judgment. With TMJ based on the concept of Experimental Questionnaire (Lind, 1982b), the presumed (hypothetical) judgment dispositions are built into the questionnaire situation as “factors” so that they can be tested using the example of an interaction between the subject (person) and the questionnaire (situation). From the characteristics of the situation considered in conjunction with the response pattern, we can infer the extent to which the presumed dispositions have become effective in this interaction. TMJs do not demand from the individual much introspection and self-description and, therefore, are not troubled with the diagnostic distortions these demands create (Shweder, 1975).

The MUT consists of two sub-texts, the MUF of six, in which decision conflicts and lines of argument are presented. In each case the dilemma and a description of its solution is followed by the question: What does the subject think about it? Then come six arguments, on which the subject indicates how acceptable or how valid the particular argument is for him. Thus, each argument exists in three meaning-contexts accounted for in the tests; they build a three-factorial, completely crossed experimental design (Dilemma x Pro-Con x Stage), that can be extended according to the research interest. In longitudinal or experimental studies, the factors Time and perhaps Treatment and their combination with the foregoing factors are also introduced.

The “Moral” Dilemma

In test application practice, the subjects first take note of the description of the objectively given situational content, including the spatio-temporal background. This information is identical for all subjects but is open to the subjective interpretation and definition by means of which the individual makes sense of the situation. only if the situational definition proceeds according to ethical criteria does the situation become morally relevant, that is, a moral dilemma (Kohlberg, 1969, pp. 393-397; Krämer-Badoni & Wakenhut, in this volume). To suggest such a situational definition or even to make one possible,
several conditions have to be fulfilled. The situation should be constructed a priori as a moral dilemma in which two or more moral values are in conflict. Kohlberg (1976) names a total of 11 “moral issues”: punishment, property, feeling, authority, right, life, freedom, equality, truth, conscience, and sexuality.

The subject should be able to put himself in the actor's place in the dilemma. As a rule, this condition can be satisfied by situations that relate to the subject's life-world and life-experience. These situations can involve dilemmas which the subject has experienced himself or dilemmas which the subject has observed or experienced vicariously through the description of others. To empirically ensure the symbolic accessibility of a situation, the question can be posed at the conclusion of a dilemma, whether or not the subject can actually place himself in the dilemma. What is essential for the dilemma's effectiveness is not the frequency in which it is experienced in everyday life but the degree to which it presents a moral challenge.

As we have seen, the level of moral reasoning is affected by the choice of the dilemma “issue(s)” (see Eckensberger & Reinshagen, 1980, p. 80). The preference hierarchy for the argumentation stages already points to the fact that not every dilemma must be solved at the highest stage, i.e., Stage 6. Whereas for some dilemmas (e.g. the Doctor Dilemma) the arguments are most strongly accepted at Stage 6, in the case of the Worker Dilemma (see the MUT in the appendix to this chapter) reasoning at a lower stage, that of Stage 5 is more strongly accepted. The latter does not necessarily imply a lower stage of judgment competence since, from a moral point of view, the workers' dilemma may indeed be most adequately dealt with on Stage 5 rather than on Stage 6. This reminds of the fact that not every dilemma is a Stage 6 dilemma, and that many dilemmas are not even moral dilemmas. Hence, a highly developed competence in moral judgment can mean to prefer a lower stage of reasoning over Stage 6. This is not paradoxon if one recognizes that dilemmas are not equal. It is, therefore, important to choose at least some of the dilemmas to be used so that they require a discussion at the postconventional stages. That is the case when human dignity or human life is made the issue. Which other issues and situations are selected for the dilemmas also depends on the study's interest. If the dependency of moral judgment on specific areas of life, the so-called segmentation effect (Döbert & Nunner-Winkler, 1975), is to be studied, then the areas of interest also have to be represented in the dilemmas; this was true in the MUF, in which some of the dilemma were written for the military life-world of soldiers fulfilling their military obligations in the armed forces, while the other dilemmas were related to their civilian life-world. The apparently plausible demand for ecologically valid situations (see Pawlik, 1976), that is the demand that the test statistically represents a particular environment, cannot be fulfilled, because the basic
Testing Moral Judgment Competence

totality of morally relevant situations in the area of life under study cannot be determined. Rather, the dilemmas (like the arguments) must be representative of the research questions and the theory. This agrees well with the “moral definition” on which the dilemma's construction has been based. The subjects have used less frequently the term “moral,” since in daily life this term often connotes “conventional” or “norm abiding.” Asked to give the kind of problem they found this Dilemma to be, first semester university students \( (N = 538) \) defined it predominantly as a “humanitarian” dilemma (Table 5). In the case of TMJs, theoretical representativity can be secured in three ways: by basing the selection of dilemmas on moral philosophy, by submitting the dilemmas to experts' ratings, and by letting the subjects define the kind of problem at stake in a dilemma. TMJs are regularly validated in the first and second way. The subjects' definition of the dilemmas has been assessed only for the mercy-killing dilemma.

The Arguments of the TMJs

The arguments which serve as test items have been designed to meet two requirements in particular. (a) To constitute a theoretically valid instrument, every argument presented on the TMJ must represent one and only one of the six stages of moral reasoning. In any case, it is useful to double check, using an expert rating, or better, inter-expert comparisons, whether the arguments presented on the TMJ truly correspond to the Kohlbergian “criterion judgments” which they are supposed to represent. Of course, this requirement is itself constrained by the current state of the cognitive developmental theory, which is still in the process of defining the stages more precisely (Colby et al. 1978; see also Eckensberger & Reinshagen, 1980; Oser, 1981b; Lempert, 1982). (b) The subject's selection of an argument should be made apart from the decision it leads to and independently of the stage of moral judgment to which it belongs. To this effect, for each of the directions a decision can take (pro and con) and for each of the six stages, at least one argument should be included in the questionnaire so that the stage of moral orientation can be logically determined independently of the subject's opinion concerning the outcome of the moral dilemma. As we will see below, this independence permits the analytical distinction between the structure and content of moral judgment and frees the hypothesis concerning the empirical correlation between these two aspects (Kohlberg's parallelism thesis) from the reproach of tautology (Phillips & Nicolayev, 1978; Eckensberger, 1984).
Sometimes it is difficult to formulate meaningful arguments for all six stages of moral judgment. “Meaningful” means here that the arguments should not seem ridiculously or artificial at any stage. The empirical findings indicate the success of such constructions. The pro arguments are ranked in the same order by the proponents of the action described, as the corresponding con arguments are rated by the opponents. That does not exclude the fact that certain opinions, e.g., favoring racial discrimination, are inconceivable on principle at a postconventional level (Kohlberg, 1976), from our own experience we know the difficulty involved in formulating “good” reasons at Stage 1 or 2 for certain issues (e.g., for euthanasia). It is still an open question whether these difficulties are a matter of principle or resolvable only by a thorough familiarization with the methodology of the cognitive development theory.

**Strength and direction of Moral Attitudes**

Table 2

<table>
<thead>
<tr>
<th>Definitions of the Mercy Killing Dilemma</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As a humanitarian problem</td>
<td>51.3%</td>
</tr>
<tr>
<td>As a moral problem</td>
<td>25.8%</td>
</tr>
<tr>
<td>As a religious problem</td>
<td>10.6%</td>
</tr>
<tr>
<td>As a legal problem</td>
<td>6.5%</td>
</tr>
<tr>
<td>As a social problem</td>
<td>5.0%</td>
</tr>
<tr>
<td>As a scientific problem</td>
<td>0.7%</td>
</tr>
</tbody>
</table>

*Source: University Socialization Research Project, University of Konstanz.*

In the TMJs a numerical reaction scale is given for each of the arguments, on which the subjects state, in the MUT, how acceptable the argument appears (from -4, completely unacceptable, to +4, completely acceptable) to them or, in the MUF, how well one can justify or reject the action described by this argument. Giving possibilities of graded reaction permits the direct study of degrees of affective tendency, which in Kohlberg’s interviews are only inferred from response frequency. The number of choices in the scale may be adjusted for the population being studied. For persons of a higher intellectual level nine choices are provided for, which lowers the chance of ties among the item ranks. For other persons, three to five reaction possibilities are probably more appropriate (see Rohrmann, 1978). In any case, answering can be facilitated if the extreme reaction possibilities, or even all of them, are paraphrased from the numerical scale. Furthermore, it is important that there be an odd number of possibilities: subjects should not be forced in one judgment direction by even-numbered response scales in which the middle category is missing.
Finally, to deal with the outcome of a moral dilemma itself, the so-called opinion question is posed after the dilemma but before the arguments. The subject's answer to this question determines whether the pro and con arguments are to be classified in his or her case as “conforming to opinion” or “contrary to opinion.”

Three-Step Strategy for the Evaluation of the Tests for Moral Judgment

Our experimental tests for moral judgment competence are set up so as to allow measurement of the two principles, related but not identical aspects of moral judgment behavior: the content-affective aspect and the structural-cognitive aspect. In classical attitude tests only the first of these two aspects is aimed at. In the Kohlberg interview, the structural aspect is supposed to be dealt with – as far as possible – free of content. Actually, in this case the measuring scores represent a mixture of the two aspects, namely the stage a person prefers in his or her reasoning and the consistency of this preference. At first we also intended that our TMJs resemble Kohlberg's interview method by having the various stages scored with a single, possible structural value. Corresponding structural evaluation strategies are presented below. Efforts in the last few years have shown not only that results comparable to those of the Kohlberg interview can be attained (see Krämer-Badoni & Wakenhut in this volume), but also that the same problems are encountered. In order to distinguish clearly the two above-mentioned aspects of moral judgment, we found it necessary to modify the methodology of the cognitive development theory. Instead of assigning only one stage to describe the cognitive-moral development of a person, we had to represent two developmental dimensions simultaneously, namely, (a) the stage of moral affects, i.e., the highest moral stage preferred, and (b) the level of moral cognition, i.e., the stage of morality which one consistently applies in one's judgment behavior in an integrated and differentiated manner. Neither dimension can be reduced to the other, nor can one be fully understood and valued without the other. Mere preference for moral maxims says nothing about competence in moral judgment, because the determination of that sort of ability cannot be specified until the subjective moral reference is taken into account and is put into the place of the alleged objectivity of others (on the problem of “objectivity,” see Adorno, 1980, p. 84). By dealing with the content and structure of moral competence reciprocally, this cognitive moral psychology is distinguished from approaches that deal with moral behavior using the conventional definitions of what is moral.

To avoid separating the two aspects of moral judgment, the methodology
used in our TMJs is directed at the question of whether or not an individual is capable of making judgments based on his own maxims. Measurement in a TMJ makes assigning a score to the direction and intensity of a moral attitude dependent on the dispositions actually operating in a person's moral judgment behavior; furthermore, it makes recognizing cognitive (ability) components dependent on behavioral consistency relative to individually accepted categories.

**Definition of Cognitive-Structural Scores**

We view the relationship between moral maxims and behavior in the individual person as an important as well as extremely interesting object of psychological research. We cannot assume a priori that moral maxims have absolutely no influence on concrete judgment behavior in every case. Tests dealing with competence in moral judgment should be especially sensitive to this point. We have assumed that the question of whether, and to what extent, moral orientation is cognitively anchored is a serious question, and one which is capable of being answered. But the development of this question cannot be processed by the traditional methods involved in the construction and evaluation of questionnaires (cf. Lind, ch. 8, in this volume). In this context it is important to remember that cognitive moral development is characterized not only by a great cognitive anchoring of moral maxims in action, but also by the necessary process of breaking up old behavioral ties (differentiation) and taking situational circumstances and competitive values (hierarchical integration) increasingly into account. Without referring definitely to contents in interpreting behavioral consistency, there can be no distinction made between progressive integration and regressive rigidity, nor between a developed differentiation and a retarded lack of orientation. Research on “pure” cognitive structures and styles would not provide the means necessary to distinguish between these opposites. “Consistency” is a bivalent relation whose univalent counterpart (“consistency per se”) is incomplete and leads necessarily to ambiguities in psychological measurement.

Defining response consistency in relation to motivational contents in the design of experimental questionnaires enables us to define various measures of cognitive structuredness unequivocally in relation to various action-guiding dispositions. Thus with the MUT or the MUF, degrees of orientation for a person's judgment behavior can be ascertained with regard to (a) one or several of the six moral types described by Kohlberg (we call this moral structuredness, or the Stage factor), (b) the subject's opinion about the solution of the particular dilemma which he or she has settled on before the discussion of arguments (opinion conformity, or Pro-con factor, and (c) the specific dilemma context (issue differentiation, or the factor of interaction between Stage and Dilemma).
By combining the orientations (a), (b), and (c) hierarchically, one can formulate and empirically test a number of more highly differentiated hypotheses concerning the formation and function of moral judgment competence and its development, as we shall see below with the help of a concrete example. Multivariate analysis of variances (or sum of squares) is the technique best suited for calculating the scores to be assigned to the structure of the subject's moral judgment. The design variables (Dilemma, Stage of argumentation, Pro-Con or opinion conformity), or more exactly the subject's representations of these variables, comprise the hypothetical determining factors of judgment; to the degree to which these account for variance in judgment behavior, they are considered the manifest determining factors. Thus, the dependent variable is the total pattern of an individual's acts of judgment.

**Definition of Affective-Content Scores**

Once it has been established that moral categories do play a role in individual judgment behavior, one can proceed to assess the direction and intensity of the subject's preference for any stage of moral argumentation. The dual aspect concept in the Experimental Questionnaire method proves its worth here because the ambiguity of interpreting the intensity index can be reduced – which is not possible in classical attitude measurements (Nunner-Winkler, 1978). This is because from the structure analysis in a TMJ one can distinguish between intensity and consistency of an individual's judgment behavior. So, for example, a comparison of preferences for the six moral stages lets one decide between the two interpretations. If there is variation among an individual's judgments but no significant differentiation in regard to moral concerns, then the degree of intensity has to be interpreted as an expression of response behavior that is “inconsistent in relation to moral categories.” Apart from this, behavior can indeed be consistent in relation to other orientation categories, e.g., in relation to the subject's attitude to the whole questionnaire (acquiescence) or in relation to the arguments' conformity to one's opinion.

According to the usual standards, the influence of a variable on another variable is said to be “significant” if their correlation attains a certain magnitude depending on the number of measurements. In the MUT, which contains 24 items, any correlation between the stage factor and an individual's judgments larger than $r = .33$ (which is equivalent to 10 percent of the variance of judgment) is significant on a level of $p = 0.05$. In the MUF this requirement is fulfilled with five percent variance since there are four dilemmas or 48 items. If one believes, as does Mischel (1968), that such low coefficients represent a “barrier” in personality research, then with this criterion one could fear that the effectiveness of a moral disposition can hardly be proven.
Fortunately, this impression which is derived from personality research using self-descriptive questionnaires, is not supported by research with TMJs. The average range of common variance lies far above the 10 percent barrier. In the studies to date, the average acceptability of stages or levels or their rank (modal stage or modal level) were meaningful in most cases and could be properly calculated. As very different studies have shown, the rank order of stages which is the main index of the affective-content aspect of moral judgment differentiates only slightly between subjects. A preference for the higher stages of moral judgment can already be found among young people and children; even across various cultures there seems to be a consensus in regard to this rank order of moral stages (see Damon, 1977; Piaget, 1973/1932; Rest et al., 1969).

For a differential psychologist to be concerned with this affective aspect seems inappropriate, insofar as his research is limited to differences between people. Psychology must not, however, neglect the intra-individual dimension especially its affective aspect. It is of great importance for (a) legitimizing the competence concept (in any case it relieves the psychometrician to know that his stage conception is also shared by the person affected), (b) clarifying the question of what initiates cognitive-moral development (moral affects represents, among other things, a fact to which the social agents can appeal in case of a controversy), and (c) analyzing of situational differentiation and segmentation.

Definition of a “Stage” of Moral Development

Although the concept of stages of moral development is controversial, the simplicity of such a concept and its fruitfulness in the interview studies conducted by Kohlberg and others make it seem desirable to construct appropriate indices for closed tests, too. Prerequisite to that, normatively and methodically, is that the orientations to stages also be established empirically as a developmental order, i.e., that (a) preferences for stages form the ranking theoretically expected, and (b) this aspect concerning content parallels the cognitive-structural aspect.

Both assumptions could be empirically confirmed. (a) In all the studies which have used the MUT and MUF, the theoretical preference ranking was confirmed. Even though single individuals and single sub-tests show deviations, on the average there emerges as a social fact a consistent pattern of preferences which increase from Stage 1 to 6 (see Lind, 1978a, 1984a, 1984b). (b) Correspondingly, we have found the second assumption to be empirically confirmed with similar clarity. In all studies, a pronounced pattern of correlations between the cognitive and affective components of moral judgment always
emerged, which was in perfect conformity with the theory (Lind, 1978a, p. 190, 1984a).

However, although there are a number of very suitable procedures aimed at measuring the cognitive stages of moral development, the task of assessing them is not yet complete. These procedures share the common intention of dealing with the moral orientation which is most highly preferred and cognitively anchored in an individual's acts and thoughts. This means that an individual's judgment is consistent in the area between Stage 1 and the highest stage that is cognitively established, and is inconsistent in the area that lies "above it." For groups that show a relatively homogenous development, such a partial consistency among individuals is easily demonstrated. For German high-school graduates and beginning university students, three kinds of analysis (longitudinal, correlational, and community analyses) point quite clearly to a shift in consistency that takes place between Stages 3 and 4; for university graduates the shift occurs between Stages 4 and 5 (Lind, 1980d). This finding points to the fact that moral judgment competence in the first group is cognitively established until Stage 3, and in the second group until Stage 4, a finding which is consistent with the more recent results by Kohlberg and his associates (Kohlberg, 1973, 1979a).

The (intra-)individual analyses, for which the tests dealing with moral judgment competence are actually designed, use various measurements for individual response consistency in order to ascertain cognitive stage scores for single individuals. In the “Entwicklungslógische Skalievectoren” method (ESV) (stage vectors logical to development), ideal types of response patterns for each stage are compared with real ones (Lind et al., 1976). Wakenhut (1982) has extended this method by utilizing the phenomenon of partial consistency for scaling an individual's moral judgment. His computer program MODI assigns a stage of cognitive-moral development on the basis of an empirical consistency test of preference values. Briechele (1981) has suggested a similar but simpler evaluation algorithm in which the cognitive stage attained in every dilemma is by definition the lowest of those stages that have been preferred by the subject once and rated at least neutrally one other time.

As another interesting form of evaluation, Heidbrink (in this volume) discusses the possibility of calculating intra-individual correlations for stage preferences with the theoretical stage sequence (from Stage 1 to 6) using Kendall's $\tau_a$, and to use these as a measure for moral judgment ability. But also the widely used individual score reference profiles permit a differentiated analysis of the preference hierarchy for the stages (content) and their cognitive establishment (structure), and thus have proven their value to research (see Rest, 1979a; Lind, Sandberger & Bargel, in this volume; Oser, 1981b).

83
Anatomy of a Sample Moral Judgment Behavior

The differences and similarities among the previously discussed measures of content and structure of individual moral judgment competence can be illustrated by the response patterns of a concrete albeit hypothetical individual, to the MUF, as portrayed in Table 2. We will assume that the subject is able to imagine himself equally well in all the dilemmas, such that no dilemma need be excluded from the evaluation.

(1) Using the individual analysis of variance, the analysis of this sample shows that in his judgment behavior the subject orients himself slightly, but still noticeably, toward moral categories. The degree of determination by the Stage factor is 14.8%. What is especially noteworthy here, however, is that the degree to which the subject's judgment behavior is determined by the Pro-Con factor (34.5%) greatly outweighs his concern with moral stages. Further analysis shows that the combination of Stage and Pro-Con factors, that is, the acceptance of reasons both on grounds of their moral quality and their opinion conformity, is also of great importance for this person.

(2) The analysis of the content aspect of judgment competency results in a modal preference for Stage 4, as portrayed in Table 2. In the present case this applies to the response matrix with a conforming opinion as well as a response matrix with an opposing opinion (Table 3). For a rougher characterization, it would suffice to calculate the modal level, which in the present example is the conventional level (B).

(3) To ascertain the subject's stage of morality, a mere preference stage, as we have seen, is not at all sufficient. one can also not expect at the present stage of discussion that the various procedures common until now converge on one value. In our example, the calculation of a stage score with the help of the MODI program shows Stage 4, using Briechle's calculation, Stage 3.

Finally, out of the multitude of other, quite diverse topics of evaluation by TMJ (see Lind, 1978a; 1980b; Wakenhut, 1981a), the phenomenon of context-dependent segmentation should be mentioned here. Segmentation becomes apparent if the initial matrix (portrayed in Table 2) concerning the subjects opinion about the dilemma is broken down into an opinion-conforming and an opinion-opposing response matrix (Table 3). As the varying mean values in
the example indicate, it is generally more difficult for this subject to prefer arguments that are directed against the opinion which he has already established as his own. If, following Heidbrink, one calculates separately the structure values for the opinion-conforming response profiles \((\tau = 0.46)\) and opinion-opposing \((\tau = 0.55, \text{ both corrected for ties})\), then it becomes clear from the discrepancy that particularly the latter points to a low degree of cognitive establishment of moral maxims. Commital to an opinion about a concrete dilemma before judging arguments represents an important “situational threshold,” on which moral competence has to stand the test, as we saw earlier in this chapter.

In the response to opposing reasons, the arguments preferred for both of the dilemmas form the military life-world – the Schneider and Neumann dilemmas – are consistently lower after Stage 3 compared to the arguments preferred for the other two dilemmas, which are from the civilian life-world. That is contrary arguments at conventional and postconventional stages are less often preferred or allowed in a military context; military and non-military life-worlds are possibly moralized in different ways by this person. This initial reference to the segmentation of one's moral judgment can be further pursued in a two-way analysis of variance of the responses to opposing reasons. The first factor is defined by the moral stages, the second factor by the dilemma included in the two contexts (military vs. non-military). The degree determination...
of the individual judgments by Stage factor and the Pro-Con factor (dilemma context) combined is, at 34.2%, in fact significantly higher than that by the Stage factor (7.0%). That is, this person differentiates among moral reasons only if they are conforming with his opinion. Opposing reasons are differentiated to a lesser degree.

Concluding Remarks and Open Questions

Structural tests dealing with moral judgment competence were designed to show that various possibilities for structural evaluation can be directly related to the cognitive theory of moral development. That is exactly what distinguishes this approach from other psychological conceptualizations; for example, in the deservedly criticized concept of attitude there is a gap, difficult if not impossible to close, between the theory and the empirical-operational access to it. Furthermore, the idionomic scaling method characteristic of our TMJs is independent of the distributions of traits found in particular samples. The test for consistency in moral reasoning is conducted on an *intra*-individual and not an *inter*-individual basis. In this way our approach takes into account the numerous and well-warranted criticisms of classical attitude and personality measurement in general and its application to moral psychology in particular (Kohlberg, 1958, 1979; Broughton, 1978; Eckensberger et al., 1980). The ideas discussed here can also be applied to approaches outside the tradition in cognitive-developmental theory. Potential areas of application include those of cognitive-social development (Selman, 1976; Edelstein & Keller, 1982), ego development (Loevinger, 1976), the development of interaction (Oser, 1981b) and other areas in which a comparable developmental logic has been formulated. The TMJs were already used in a number of studies to examine the conditions, course and results of socialization processes. In those studies, not only was the cognitive theory of moral development confirmed on a far-reaching scale (convergent validation), but also new insights were made possible into the way the content and structure of moral competence are related (see Lind, 1985d, 1985e).

However, the discussion of these test methods continues. There are a number of open questions and problems of a methodological nature that confront the findings, but confirm the fertility of this approach, and which are not confined to the tests presented here. According to the cognitive theory of development, it would be inappropriate to evaluate these methods using the criteria of classical test construction, reliability and validity, as is sometimes done in criticisms of research on the cognitive theory of development (see Kurtines & Greif, 1974; Eckensberger et al., 1980; Schmied, 1981). Thus, within the cognitive-structural paradigm – and perhaps even within that of classical psycho-
metrics, as Lumsden (1976) argues – it is not meaningful to determine “reliability” for a TMJ in the classical manner. The hypothesis that a mediating random process takes place between “true” moral orientation and the measuring value has been proposed often enough, but has seldom been tested empirically. It therefore seems reasonable, as Kohlberg (1979) has emphasized in his discussion of psychometric approaches, to start with a “manifest construct” model in which inconsistencies can be analyzed systematically. Structural wholeness cannot be equated with test reliability, but rather refers to “a developmental task that can only be completely carried out at the highest stage” (Lempert, 1982, p. 117).

The criterion of “validity” will be treated in a similar way. To speak of the validity of a test (“How well does this test measure that which it is supposed to measure?”) assumes the existence of an elaborated theory upon which the test is based (“How well does the researcher know that which he wishes to measure?”). The tests presented here are, in regard to several central hypotheses, capable of demonstrating the empirical validity of the cognitive-developmental theory. That capability leads to the further conclusion that the tests themselves are to some extent theoretically valid. Nonetheless, it may well be that one should possibly still argue, in the manner of Broughton (1978), as follows: on the one hand, to test theoretical positions one must first assume that the theoretically established test is valid and that it is the responsibility of the theory to account for its predictions. On the other hand, if the theory is considered to be sufficiently supported, then one must question the way the theory has been operationalized. Both critiques cannot be carried out simultaneously, but they can be managed by means of an interactive procedure in research which Kohlberg and other have called “bootstrapping.” In sum, it seems clear that the validity of measuring methods is embedded in the validity of the research processes in which they are used, and therefore that the development of our instruments is closely bound up with progress in theory foundation.

This chapter is a revised version of the original German chapter. It also contains, especially in the discussion of the evaluation of the TMJs, many alterations of our original statement (Lind & Wakenhut, 1980), made in the light of recent critical and constructive suggestions on the MUT and MUF (Eckensberger et al., 1980) as well as of the results of our recent research.
Notes

2. The concept of “competence,” which has many meanings, is chosen to emphasize the aspect of cognitive ability involved in every moral judgment. It is preferred to the concept of “consciousness” because we, like Habermas (1976a, p. 86), do not want to confine our analysis to conscious judgment behavior. Our use of the concept of competence differs from that of other authors, who equate competence with a value posture that is not yet translated into action (performance). On the contrary, competence here means precisely the ability to translate moral attitudes into specific judgment behavior.
3. For other methods of assessing moral attitudes without reference to cognitive aspects, see Pittel & Mendelsohn, 1966; see also Kuhmerker et al., 1981; Eckensberger et al., 1980; Rest, 1973.
4. This research is supported by the Deutsche Forschungsgemeinschaft (German Research Council). The MUT has been developed in close collaboration with the following members of the research group: Tino Bargel, Barbara Dippelhofer-Stiem, Gerhild Framhein, Hansgert Peisert (director), Johann-Ulrich Sandberger, and Hans Gerhard Walter.
5. The research has been conducted at the Sozialwissenschaftliche Institut der Bundeswehr (German Armed Forces Institute for Social Research), with Thomas Krämer-Badoni (now at the University of Bremen), Eckehard Lippert, Paul Schneider (Staatsinstitut für Schulpädagogik, Munich), and Roland Wakenhut (University of Augsburg).
6. For further instruments developed on the basis of the MUT see Briechle, 1981; Hinder & Kanig, 1981.
7. It is easily overlooked that in mathematical statistics “error” is defined precisely as the result of a pure random process. This means that “error” is a particular fact which, like any fact, can be investigated empirically.
8. For a discussion of “hypothetical” versus “real” dilemmas, see Damon, 1977; see also Döbert & Nunner-Winkler as well as Oser & Schläffi, both in this volume.
9. Kohlberg and others sometimes call the opinion concerning the solution of dilemmas (“What should X do?”) the “content” of moral judgment. But we have already used this term to designate the affective aspect of moral judgment, so in the present context we have decided to use the term “opinion” in order to avoid ambiguity.
10. Concerning the problem of “ideological bias” of moral psychology, see Pittel & Mendelsohn, 1966; Simpson, 1976; Hartmann, in this volume.
11. The Dilemma factor by itself does not indicate, as we once believed (Lind & Wakenhut, 1980, p. 322), the phenomenon of segmentation.”

In meanwhile (2016), the MJT (German: MUT) has been renamed to Moral Competence Test (MCT). See G. Lind (2016). How to Teach Morality. Chp. 4. (Berlin, Logos)
There is a gap between the expectations concerning the socialization effects of university and the respective empirical findings. On the basis of common sense, many people would expect a university or college education to “make a difference,” not only in regard to specialized vocational skills but also generally, in regard to such supra-vocational “skills” as critical thinking, judgment ability, and social responsibility. Contrary to this normative expectation, however, some researchers have concluded that university socialization has no effect on such skills, or at least no general and lasting effect.¹

Two basically different explanations for this gap are suggested: Either university education actually fails to reach its proclaimed aims, or the “university-makes-no-difference” finding is due to our inability to actually “see” those supra-vocational effects, i.e., it results from shortcomings of the concepts and instruments employed in the studies on which that finding is based. The first explanation can be contested only indirectly, e.g., by providing a measurement methodology which improves our ability to assess the effects of socialization. If the results remain the same, we would then have to accept as a matter of fact that university education fails to achieve its aims. However, if we could demonstrate a general democratizing effect of university education by using better methods, we could refute the implicit assumptions of the research methods which have produced these “no-difference” findings.

In this paper I shall try to show that, indeed, the concepts of attitude and attitude change, which were the core concepts of the impact of college research, have been defined too narrowly to adequately assess such effects of university education. Without a doubt, the development of classical attitude tests represented a major step forward in social research (see, for example, Feldman and Newcomb, 1970; Cloetta, 1975), primarily because, in comparison to case study methods, such tests are more transparent and more applicable to large scale surveys. This implies that they are objective and can easily be criticized and improved. However, the classical concepts of attitude and attitude change limit the possible outcomes of socialization in two ways: (1) in regard to the evaluative aspect of attitudes, and (2) in regard to those attitudes which differentiate among persons. In doing so, classical attitude research ignores a central aspect of educational outcomes, namely, the cognitive aspect of attitudes and its structural transformation. Consequently, this approach re-
duces the process of socialization to changes in affective magnitudes. This self-confinement also seems responsible for the fact that the process of socialization is often perceived only under the narrow categories of “adaptation” and “deviation.”

The cognitive-developmental approach (cf. Kohlberg, 1973a; Perry, 1970) offers a conceptual and methodological alternative for higher education research. From this point of view, the general democratizing effect of university education is conceived of, and assessed as, the development of both affective and cognitive aspects of personality. Accordingly, the effects of socialization are not reduced to changes in affective magnitudes but are construed as more complex processes of integration and differentiation. By viewing the process of socialization through the wider conceptual “lens” of cognitive developmental theory, we are able to “see” effects of university education which have hitherto been invisible to the researcher. From this point of view, even using classical attitude research, we have obtained indications of a sequential, irreversible development of supra-vocational competencies in university students (see also Lind, 1985b). These show that, contrary to the findings of classical attitude research, university education in fact facilitates, or even stimulates, the development of moral and democratic competencies among students in general.

Socialization as Attitude Change

*Attitudes in Impact-of-College Research*

Until recently, the concept of attitude has predominated in research on the effects of university socialization (cf. Jacob, 1957; Newcomb, 1957; Sanford, 1962; Feldman & Newcomb, 1970; Dressel & Mayhew, 1971; Lenning et al., 1974; Cloetta, 1975; Dann et al., 1978; Huber & Vogel, 1984). This proposition is not invalidated by the fact that only a few studies contain an explicit definition of their research subject. The theoretical assumptions of these studies can be perceived in their concrete research methods which define in operational terms the constructs being employed. There are many variants in research methods. However, there are some core features of attitude testing which are common to nearly all studies and which are based on classical testing theory. This makes it possible to speak of a unique paradigm: the “classical attitude concept” (cf. Scott, 1968; Fishbein & Ajzen, 1975). The concept and its measurement contain, as we will see below, certain psychological assumptions about the nature of the human mind. Thus, the central question is whether these assumptions are compatible with the research hypotheses which they are being used to test.
Hypotheses of Attitude Research

One major focus of research into university socialization deals with its hypothesized democratizing effect, i.e., with the democratic personality it is supposed to shape. This includes attributes such as innovative competence (liberalism), critical judgment ability, moral autonomy, willingness to assume social responsibility, and general liberal attitudes. Of particular interest in research has been the dimension of "conservatism," which is considered a kind of antipode to the democratic attitudes (liberalism) expected of university and college graduates (cf. Newcomb, 1957; Cloetta, 1975). The hypothesis to be tested is: Is the university capable of instilling democratic attitudes in the student? Or is it appropriate to resign ourselves to the conclusion that the university has no, or no lasting, influence on attitude change?

Results of Attitude Research

Until now the research results have seemed surprisingly clear – and negative. When asked what college does for the individual, the college researcher Theodore Newcomb (1974) answered: "Frankly, very little that is demonstrable" (p. 73). The findings of impact-of-college research confirm this conclusion in that they are concerned with the prediction of the non-vocational effects of college education. With only a few exceptions, research either does not show any of the anticipated attitude changes, or it shows that such changes are revised at the end of college or in the initial phase of the individuals' professional life. Even Jacob (1957), in his highly regarded summary of a quarter-century of attitude research, ascertained that college had no significant impact on attitude change. Feldman and Newcomb (1970), after studying extensive research, concluded that college effects very little change in attitudes and values, or if there is a change, it occurs in a very specific manner. They reported varying outcomes, depending on the college attended, subject of study, and student characteristics, but one can hardly speak of a general, encompassing effect produced by college education. If one considers in particular the democratizing effect, some consistent changes were observed in the first years of study. According to Feldman and Newcomb (1970) and other surveys in this field, there is a slight but general turning away from conservative, authoritarian, non-democratic attitudes during the college years.² This trend could also be shown in a longitudinal study carried out by the research project "Teachers' Attitudes" (Cloetta, 1975; Dann et al., 1978). In this excellently designed research, many new insights into the process and conditions of socialization could be gained, demonstrating that, at least in regard to profession-related attitudes, higher education has some general and stable effects. But in regard to
more general conservative attitudes, the initial trend in the liberal direction is reversed. The democratizing effect abates again at the end of college and in the initial phase of work. The authors refer to a “practice shock,” which cancels the effect of college.³ This seems to prove that students' attitudes do not, as Newcomb (1974) had expected, stabilize, but instead adapt to the particular climate of opinion in the environment: i.e., attitudes are ephemeral and fleeting phenomena in the course of development.

Problems of Interpretation

Must we therefore repudiate the idea that college graduates gain supra-vocational abilities? Are the findings so evident that one can consider colleges and universities ineffective in fostering democratic competencies? We will see that these questions cannot be answered with an unambiguous “yes.” First of all, these findings are influenced by the methods used and hence are theory-imregnated interpretations which are debatable. Of course, these findings are not produced completely independent of reality, but are constrained by the methods of assessment and data evaluation. Therefore, the fact that no effect of higher education was ascertained does not necessarily mean that no effect exists.

Indeed the findings of attitude research give rise to several problems of interpretation. Concept and measurement of attitudes have long been criticized (cf. Converse, 1970; Deutscher, 1973; Meinefeld, 1977). There have been attempts to discover a satisfactory explanation for empirical findings by introducing auxiliary assumptions without abandoning the classical attitude paradigm. As far as I can see, however, most of these attempts have created new, unsolved problems. For example, to solve the reliability-change dilemma, it was suggested that we require not only the stability of attitude scores (more exactly: their rank order for all persons in a group studied), but also the stability of change rates (i.e., to determine “reliable” types of changes). However, since these two demands contradict each other, attitude scales that meet the new criterion would have to be eliminated according to the old criterion, and vice versa. This dilemma is especially evident in the conservatism scores which have a comparatively high level of rank reliability, but whose reliability of change is, on the average, close to zero (Hohner & Dann, 1978; see Figure 3 below).⁴

Equally paradoxical are practical suggestions associated with this interpretative framework. If we retains the attitude paradigm, we have to demand that the socialized person adapts to democratic values and at the same time resist them. Not surprisingly, many authors evade normative questions and are reluctant to determine precisely the direction and intensity of the attitudes most
compatible with the goal of socialization. To understand these paradoxes and ambiguities, it seems necessary to revisit the classical concept and method of attitude research. (For a discussion of these problems of interpretation and valuation, see Cloetta, 1975, pp. 180ff.; Dann et al., 1978, p. 90; Fend, 1971).

Critique of Attitude Concept and Measurement

It is surprising that many authors neglect to define explicitly the term “attitude” in order to go on to questions concerning the origin of, and changes in, attitudes. Feldman and Newcomb (1970) discuss the concept of attitude only tangentially, although their reports almost exclusively contain studies based on the classical attitude paradigm. Some, like the authors of the “Teachers’ Attitudes” study who refer to Kelman’s (1974) definition of attitudes, outline a complex conceptual framework which, however, may not be fully compatible with the classical research methodology they actually use. The closest we have come to a definition has been in methodological treatments of attitude measurement (cf. Thurstone, 1931; Scott, 1968; Dawes, 1972). Of course, almost all the studies use the same kind of attitude tests and hence the same implicit definition of attitudes. This more or less implicit definition and its limitations – but not the intentions and allegations that accompany them – will be examined here.

An “attitude” is defined as the “degree of positive or negative affect associated with some psychological object” (Thurstone, 1931). This definition is the starting point and a basic part of most methods of attitude measurement (Scott, 1968), even if some researchers do not go beyond it to include purely descriptive statements, as well as affective ones, in their attitude scales. Thus, in classical terms attitudes are distinguished solely by their direction and intensity. This creates a problem: if attitudes must be either positive or negative, no conceptual tool is available for distinguishing among average, neutral, or other cognitively differentiated attitudes. When a person gets a medium score on an attitude scale, its significance must remain ambiguous in that such a score can mean that he or she has either (a) no attitude, (b) a conflicting attitude, or (c) a highly differentiated attitude (cf. Shaw & Wright, 1967, pp. 7ff; Converse, 1970). Neutral answers are considered uninteresting or expressive of a tendency on the part of the subject to hide his or her attitude. Hence, a researcher working within the framework of this paradigm must attempt to force the subjects to express an attitude through the use of “forced choice items.”

Furthermore, in attitude testing it is assumed that the affective tendencies under investigation are in one way or another characteristic of everyone. Classical attitude measurement is based on tests of groups of people, not of indivi-
duals. A “reliable” and “valid” attitude measurement is spoken of only when the subjects of a particular group under study meet two requirements. First of all, the subjects must differ considerably from one another regarding the direction and intensity of this attitude: i.e., “the object of attitude must be controversial in the investigated group” (Cloetta, 1975, p. 37). At the same time, however, the subjects must be similar to one another in regard to their attitude structure, i.e., the particular attitude must be present in all persons and must determine the same set of responses to the same degree. A violation of any of these assumptions leads to charges that the research instrument is “unreliable” and “invalid.” Thus paradoxically, in the moment in which the group's attitudes become either similar in direction and intensity or structurally heterogeneous, the particular “attitude” disappears or becomes inaccessible. Attitudes common to all members of a group and attitudes associated with individually varying forms of cognitive structure are eliminated from the analysis, supposedly on the grounds of purely methodological criteria, even though these attitudes are essential to the socialization process (cf. Cloetta, 1975, p. 39; Geulen, 1975, p. 86; Portele, 1975b, p. 98).

However, in focusing on the affective components of attitudes – i.e., on their direction and intensity – psychology limits itself to cases in which the attitude in question is already completely integrated but is at the same time not yet differentiated according to higher values and ideas. This occurs rarely and, is therefore of only limited interest. If taken seriously, these assumptions would prevent all attitudes from being tested, since there is virtually no instrument of attitude measurement that achieves “reliability” or “validity” indexes which agree with the basic model. Moreover, because they are restricted to affective aspects, the measured effects of socialization are limited to simple, one-dimensional changes in the affective components of attitudes. Even in multi-component theories of attitude (see, e.g., Rosenberg & Hovland, 1963), the structural aspect of the individual behavior is eliminated and only structural properties of the correlations among persons are considered. This elimination reduces the process of university socialization to a mere contrast between conformity with given norms and deviation from them. It conflicts with the notion of socialization as an

“active process, [which] differs from that widespread concept that uses as its basic terms the opposites “adaptation” and “deviation”[…]. [Because], although it is stressed in the theoretical development of these terms, its concern is with descriptive categories free of values, “deviation” is regarded as socialization which has failed and thus is, subsequently, once again excluded from the concept.” (Zentrum I, 1973, p. 852)

What can be done? If one keeps the attitude paradigm, one could simply “admit nonconformist behavior [as] a necessary element of socialization” (Zen-
trum I, 1973, p. 852). That would necessarily lead to successful socialization being defined both as adaptation and as deviation, and would thus broaden the realm of desirable educational effects so much that only trivial expectations could be derived from it, because they would be fulfilled from the outset. However, if one defines the area of desirable socialization results as the narrow, almost imaginary margin between conformity and autonomy, then this margin becomes, as Fend (1971, p. 39) pointed out, “very narrow,” possibly too narrow to be examined empirically at all.

The dilemma of socialization research is thus laid bare: things that we can study with the concept of attitude allow socialization theory to formulate only empirically empty hypotheses, and the things we really want to examine lie outside its conceptualization. Thus, the simplicity and explicitness of the attitude concept indeed “render its inadequacies obvious” (Scott, 1968, p. 208). Its major inadequacy is its failure to account for the structural aspect of attitudes in addition to the affective aspect. This should imply that neither the behavioral structure of a particular group of subjects nor a specific kind of behavioral content (so-called “cognitive beliefs”) is separate from affective content. Rather, the structural aspect refers to the relational properties of an individual’s responses, properties which are ontologically inseparable from the affective aspects (Lind, 1985a).

Socialization as Moral-Cognitive Development

With their methodical elaboration of the concept of cognitive-moral development, Kohlberg, Perry, and others have introduced a new approach to research on higher education, which has acted as an alternative to the classical paradigm of attitude change. The volume The Modern American College that appeared in 1981 follows exclusively the concept of development. According to the words of its editor, W. Chickering, the development concept should “offer a consistent solution on the basis of the unifying idea of adult development.”

However, upon closer scrutiny, it is evident that in the concept of development very different and partly contradictory ideas have merged. In fact, the concept of attitude change has been revived, as has the old concept of genetically controlled maturation. Thus socialization in college is sometimes regarded as nothing more than an episode in an individual life-cycle (Chickering, 1981b; Chickering & Havighurst, 1981). The effect of higher education is regarded as ephemeral in comparison to the effect of genetic determination, even if the “hard realistic requirements for effective social contribution” (Chickering, 1981b, p. 9) are not overlooked. It is doubtful, however, whether the university can provide, simply by imposing those hard requirements, an “in-
creased coherence and an enhanced sense of community” (Chickering, 1981c, p. 773). The development of a sense of community and the competence to cope with socio-moral conflicts will not result simply from hard requirements.

Aside from its sometimes fashionable use, the term “development,” as elaborated by Piaget and Kohlberg, provides an alternative conceptual and methodological framework for the study of socialization processes in the following respects:

**Individuality.** The subject of the analysis of moral competence is above all the individual, as well as the structure of his behavior and judgment, and not a comparison of the isolated character variables of persons in a group. Thus the assessment methods which are based on this approach ensure that the traits of an individual are defined independently of other individuals in a group.

**Manifest Judgment Behavior.** The existence of moral attitudes is tested in the manifest structure of a person’s judgment behavior. Competence in moral judgment is “a construct rather than an inference, and is warranted only on the grounds of ‘intelligible’ ordering of the manifest items.” (Kohlberg, 1979, p. 14; see also Lind & Wakenhut, in this volume).

**Affect and Cognition.** In contrast to the classical theory of attitude and the multi-component theory of attitude, the cognitive theory of development insists that “a moral act or attitude cannot be defined either by purely ‘cognitive’ or by purely ‘motivational’ criteria” (Kohlberg, 1958, p. 16). Incomplete integration of behavior in a normative orientation and a highly differentiated value posture are not, as in the attitude paradigm, indiscriminately designated as “inconsistencies.” Accordingly, we do not attribute “unreliability” to the measuring instrument. Instead we understand it to be at least partly an expression of the cognitive aspect of a judgment behavior (cf. Lind, 1985a).

**Development.** With the dual concept of affective-cognitive personality traits, it is possible to conceptualize socialization within the wider framework of a developmental theory. “Development” is defined as “changes in the form of reorganization of responses over time as contrasted with the change in the strength or accuracy of the responses...Thus, the developmentalist focuses upon structural changes in the response” (Zigler, 1963, p. 345). In regard to the area of social-moral abilities, development means structural change in the individual’s moral-cognitive system.

This wider concept of structural change allows us to present the role of the university in the educational process in a more adequate way. Socialization is not, as the theories of adaptation assume, simply a change in behavior due to altered environmental conditions, but a differentiation and integration of atti-
attitudes and norms as a consequence of the “interaction” between the structure of the organism and the structure of environment” (Kohlberg, 1969, p. 348).

Review of Empirical Findings

The two above-mentioned paradigms of simple “change” and structural “development” can be compared empirically in regard to different hypotheses about the course and result of university socialization. As we have seen, the most important difference consists in the fact that, in addition to a change in direction and intensity, development also implies a transformation of an individual’s attitude system. Development is thus to be regarded as an integral, two-dimensional change that cannot be reduced to only one dimension (or to two ontologically separate dimensions) of change without the loss of essential information.

The position of development theory can be illustrated by the concept of the “conservative attitude,” which as we noted above is generally viewed as a central indicator of the democratizing effect of college education. Whereas the classical attitude theory depicts socialization one-dimensionally as the acquisition and loss of “affective quantities” in time, the developmental model also takes into account the cognitive-structural dimension. In such a two-dimensional developmental model, phases of integration and differentiation can also be distinguished, as can changes in direction and intensity. Using this conceptual framework, we may hypothesize that people's attitudes undergo a develop-
Figure 1 Change and Development: Projection of Affective-Cognitive Development onto (a) a Two-Dimensional, and (b) a Three-Dimensional Space of Change.
mental process. As long as the individual has not yet (even unconsciously) de-
veloped a concept of “conservatism,” he or she will not be able to respond con-
sistently to conservative or progressive statements.

The scientifically construed concept of “conservatism” is not yet “repre-
sented” in the individual and thus is not measurable by a conservatism scale.
Only as the person becomes more and more acquainted with this concept will
he or she respond with increasing consistency (whether positively or negativ-
ely) to conservative statements – something that we call the “phase of integra-
tion.” After this takes place, the attitude of conservatism will become measura-
ble.

In a second phase, the same statements are no longer judged solely accord-
ing to this one category, but also in regard to other judgment criteria which
compete with it or even cancel it. In this phase of attitude differentiation, we
can observe an increasing preoccupation with context, which leads to a greater
distrust of general statements (“slogans”). On the surface, i.e., in regard to the
criteria of classical attitude testing, the attitude behavior again becomes “in-
consistent” or “unreliable” and thus non-measurable. However, whereas in the
beginning inconsistency means lack of judgment competence, it can later be
evidence of a highly developed cognitive structure. This “two-sided” develop-
ment process of integration and differentiation implied by cognitive-develop-
mental theory is contrasted with the “one-sided” process of change included in
the two-dimensional process diagram of Figure 1.

We cannot yet present a crucial experiment which would clearly demonstrate
the superiority of one paradigm over the other, but we expect to obtain firm
evidence for one or the other based on findings from the longitudinal studies
which we are at present carrying out in order to clarify this question. However,
the research that has already been carried out offers important indications
which make possible an empirical contrast of the two paradigms. Three sour-
ces are at our disposal: (a) traditional attitude studies, insofar as their results
indicate that cognitive processes are involved in socialization, (b) cognitive-
developmental research by Kohlberg and his collaborators, and (c) the initial
findings of our own ongoing study.
The Cognitive Process of Integration

(1) Classical research on attitudes provides findings for a comparison only conditionally, as it usually treats behavior in which the cognitive developmental aspect is documented as a measurement error. Therefore, references can best be found in irregularities and “errors” insofar as they are treated at all. In this respect, two studies of university socialization are of particular interest: the well-known Bennington-study by Newcomb and the large-scale longitudinal study of student teachers conducted by Dann, Cloetta, and their associates. These two studies illustrate the best use of the attitude paradigm, and also, because of their concern with the overarching democratizing effects of university education, are directly comparable in content with the cognitive-moral development model.

One of the first indications that cognitive processes played a role here was that even after very elaborate measures were taken to construct “consistent” attitude scales, it was still not possible to construct a scale in which the great majority of persons consistently answered either “liberal” or “conservative.” This problem is not unusual for attitude research. However, it means that the main condition for the use of the classical notion of attitude is violated, i.e., the condition which requires that the “object of the attitude...in the persons questioned (1) is represented, and (2) is affectively cathected” (Cloetta, 1975, p. 38).

The problem becomes even clearer if one uses attitude scales to analyze the individual response pattern of those questioned. If one supposes that the questions on a scale are all “items in the same dimension” (Dann et al., 1978, p. 87), one should expect all items to be consistent, i.e., to be answered in accordance with the particular attitude; for example, for all the “conservative” persons, the “conservative” attitude statements are consistently answered positively and the “liberal” always negatively. The usual way to fulfill this expectation is to define attitude as the mean value and the deviation from it as “error probabilities” (see Lazarsfeld, 1959; Fischer, 1974), but this definition presupposes that “error” is precisely defined and does not render the concept of attitude vague and inaccessible.

To clarify this last point, we must consider the consistency with which subjects choose the “liberal” or the “conservative” side of the response to the items on a conservatism scale. This procedure seems legitimate if, as is ideally assumed, the items are exact replications of one another and, therefore, have identical means. We should note that the Likert method of scaling disregards this assumption, although it cannot eliminate it completely. In Figure 2, on the left, two ideal distributions of answer consistencies for the conservatism scale are shown: the theoretical distribution of complete consistent answers (A), and, for the sake of contrast, the pure chance distribution of consistency values [185]
(B). On the right is placed the distribution of actual consistency values from high school graduates (Abiturienten), which is taken from our own study with the Cloetta scale (C). From this we can see that before entering university the respondents were influenced to only a slight extent by the construct "conservatism" in judging the items. The assumption that this attitude is "represented" in everyone was not fulfilled. Although the judgment behavior of the subjects definitely deviated from the pure chance distribution, only a few persons were completely oriented in their response behavior toward the attitude dimension in question. Thus, most of the subjects' answers could not be understood when assessment was based on the one-dimensional attitude.

Figure 2 Comparison of Two Theoretical Distributions of Responses (Complete Consistency, Pure Randomness) with an Empirical Response Distribution.

Note. High school graduates (N = 708). Shortened version of the MK questionnaire by Cloetta (1975). The response scales contain the category "cannot say", which was not in the original. As used here, consistency only refers to the number of the pure "liberal" or pure "conservative" responses.
concept. Obviously, their response behavior was either less integrated or more strongly differentiated than was assumed. (It is important to emphasize once more that “dimension” here means an “aspect dimension” and not a separate variable, as for example in factor analysis).

The claim that unexplained differences between subjects can be attributed not only to errors in the measuring instrument but also to differences in individual cognitive structures is supported even more strongly by longitudinal studies. If the inconsistent attitude scores were really only a function of measurement error, the relationship of systematic variance to error variance would remain stable over different measuring points in time or vary only “randomly.”

A closer look at the findings of the “Teachers’ Attitudes” research project shows that not only do the mean scores on the attitude scale change, but the reliability and consistency values change, too, following a systematic trend which runs parallel to the process of socialization.
This holds true for various indicators (see Figure 3). Although the trends are not very high, due to high values at the start, the almost perfect monotonous change in every one of these (inter-individual) structural values is astounding. Clearly, these findings are at odds with the classical attitude paradigm, in which these coefficients, characteristic of the measuring instrument, would be independent of whatever development tendencies were discovered.

Hence, we cannot say of a test that it has “high reliability” if the concept of reliability is itself undermined by the empirical evidence. However, the cognitive-developmental theory provides us with an alternative paradigm, within which our findings can be interpreted to indicate structural changes within persons, resulting from the integration of the individual attitude systems.

The findings of cognitive developmental research confirm and extend these references. In their interview study of high school and college students between the ages of 14 and 26, Kohlberg and Kramer (1969) established clear evidence that “there is an adult movement toward integration...of moral thought in its application to life” (p. 118). They observed that “the major change in moral thought past high school is a significant increase or stabilization of conventional morality of a Stage 4 variety, at the expense of preconventional stages of thought” (p. 106), and that moral development in young adults “is not only toward greater consistency of moral judgment but toward greater consistency between moral judgment and moral action” (p. 107). Hence, moral development is characterized not so much by new moral attitudes exhibited in moral discussions as by a structural process of integration of reasoning and action at new stages.

Rest (1979a, 1979b) found further evidence for this. He let subjects evaluate moral arguments in order of their importance. Here the phenomenon of the cognitive integration of moral attitudes is shown even more clearly than in Kohlberg’s interviews, in which only the frequency of reasons given is counted. Whereas at a young age the highest stages of moral judgment are already preferred, a considerable part of the integration and stabilization of this attitude does not take place until late adolescence and, in particular, during college (Rest, 1973; 1979a, pp. 133-135; 1979b, part 1, pp. 31-34). Before and during college, a relatively drastic increase takes place in the consistency of postconventional judgment preferences on different moral “issues,” from 30% of postconventional answers to 50%.

Finally, preliminary findings from our research project on university socialization point in the same direction. To test the degree to which moral maxims were cognitively anchored, the subjects had to evaluate arguments for various decisions in moral dilemmas and were asked to address not only those arguments which supported their own opinion but also those that opposed it. By gi-
Attitude Change or Cognitive-Moral Development?

ving the subjects arguments based on opposing opinions, a cognitive threshold was created through which we could directly study the process of integrating moral attitudes during college (Lind & Wakenhut, in this volume). In an inspection of MUT data on the judgment behavior of students in professional schools, high schools, and universities, and that of scientists, a clear developmental trend could be shown (Lind, 1984a).

The Cognitive Process of Differentiation

(1) In attitude studies it is inherently difficult to find empirical evidence that this process of integrating democratic and moral attitudes really takes place, since the design of these studies and the methods of analysis dismiss signs of differentiation as signs of “unreliability.” Furthermore, as I have argued above, these studies lack the methodological prerequisites necessary for investigating the cognitive-structural aspects of socialization effects. Nevertheless, some unexpected evidence that this differentiation process occurs can be found in attitude studies. For instance, Newcomb (1957) notes as a methodologically irritating effect that students in higher semesters tend more and more toward a rejection of all “conservative” and “liberal” questionnaire statements, a finding which indirectly indicates greater differentiation in their attitude structure.

(2) Direct support for this conclusion of structural differentiation can be found in the study by Fishkin et al. (1973). As these authors show, growing rejection and critical evaluation of condensed political slogans (which are most readily associated with the dimension of “conservatism”) are accompanied by a higher moral-cognitive competency. This is in line with Perry’s (1970, 1981) observation that students’ cognitive growth is revealed in their deviation from black and white judgments and their differentiation and re-integration of ethical reasoning (see also Murphy & Gilligan, 1980; Kitchener & King, 1981).

(3) Despite our use of different methods in the inquiry and evaluation stages, the findings from our research project and comparative research by other authors who have used the MUT or a similar instrument also support the structural differentiation thesis. Keasey’s (1975) finding that differentiation of judgment behavior in relation to one’s own opinion is not very advanced in young school children or in people with low general education was corroborated by Briechele (1981) and Wischka (1982). In contrast, German high school graduates (Abiturienten) and, in particular, university students were found to be more reflective, more critical and rational, and less bound to an opinion once held (Lind, 1984a; Schmied, 1981). However, moral-cognitive development does not stop at this point. As Figure 4 shows, scientists (in both the so-
cial and natural sciences) differentiated their judgment of arguments according to moral criteria even more than, for example, did high school graduates. This can be seen in the differences in the slope of the preference profiles of both groups. However, in all the studies it was also clearly established that moral categories are not applied rigidly to all social situations without reflection, but that a differentiation is made between situations with varying moral contents (see Lind, 1984a).

Conclusion

The concepts of attitude and attitude change have produced in socialization research a wealth of empirical findings which cannot be neglected or regarded as irrelevant. In particular, the research project on “Teachers’ Attitudes” has produced much new information which calls into question the conclusion which Newcomb has drawn from attitude research that higher education has no general effect on the development of a democratic personality. However, as I have tried to show in this chapter, students’ attitudes and their development cannot be fully understood when we concentrate only on the affective aspect of

![Figure 4 Structure of Moral Judgment of High School Graduates and Scientists.](image)

Sources: Scientists (N = 54, solid line) from Portele (personal communication): High school graduates (N = 708, dotted line), from the Forschungsgruppe Hochschulsozialisation. In both studies an older version of the MUT was used.
attitudes, nor when we conceive their cognitive aspect as an entity that is separate from its affective content. Hitherto supra-professional outcomes of college and university socialization – such as critical rationality, sense of responsibility, and democratic orientations – have been classified by and large as “affective” goals and contrasted with the “cognitive” goals of education, such as vocational qualifications or general belief. But, as I have argued elsewhere (see ch. 2, in this volume), we do not need to separate the two basic aspects of human behavior ontologically. So-called “affective” attitudes have their own cognitive structure, which is distinguishable but not separable from their affective contents. Thus the affective and the cognitive aspects of attitudes can be researched as a whole.

I have tried to show that this conceptual problem has significant consequences for the methods and results of socialization research. Whereas on the basis of classical attitude research – which reduces attitudes to their affective aspects – no general supra-professional outcome of higher education could be detected, the more comprehensive cognitive-structural approach makes discernible the transformation of students’ personalities, as could be expected from the educational goals that the university sets for itself. The university apparently contributes its share to the evolution of the individual personality, and thus also to the evolution of social forms of existence, although this contribution may be viewed as still inadequate in view of the social problems with which the university graduates will be faced when taking over responsible positions in our society. From this finding we can conclude that even for “initial purposes” it is not advisable to confine research to “varying degrees of favorableness and unfavorableness” and thereby disregard the structure of a person’s attitudes, as certain eminent scholars have suggested (cf. Feldman & Newcomb, 1970, p. 55; Müller-Fohrbrod & Cloetta, 1975, p. 198). We would be better advised to acknowledge the fact of structural change or development, i.e., the fact that – as Feldman and Newcomb have remarked – “there are other kinds of individual change which do not fit into [this] typology. For example, one type of change is the reordering by the person of value and attitude dimensions in terms of their salience or centrality to him. Also there may be other value and personality reorientations – such as the development of new frames of reference – that are not merely shifts along a given dimension” (p. 58n). Cognitive-developmental theory and research has provided methods for assessing this reordering and reorganization of individual attitude systems in the course of the socialization process, although these methods are not as clear-cut and as suitable for large scale research as are classical attitudes tests. As our own research demonstrates the assessment of structural change is not bound exclusively to interview methods but can also make use of attitude measurement, provided that it has been revised in such a way as to render the assessment of the cognitive component possible.
Notes

1. Universities and other kinds of institutions of higher educational (Fachhochschulen, etc.) are not differentiated here because in recent years they have become very much alike. For comparison, we frequently refer to college education in the United States, where aims (general education), are more like those of German universities than are the aims of American professional schools.


3. Müller-Fohrbrodt et al., 1978. This back-and-forth movement of attitude change has been named the Konstanzer Wanne (the Constance Dip); see Cloetta, 1975, pp. 176-177; Dann et al., 1978, pp. 37, 230; Hohner & Dann, 1978. For further studies which support these findings, see Koch, 1972, p. 135.

4. On the problem of “reliability of change” see Renn, 1973; Helreich, 1977; Dann et al., 1978, p. 221. It seems that in the past the psychometric and the statistical sides of the problem have not been sufficiently differentiated. “Regrettably, psychometric properties, namely reliability, have been the predominant concern in the behavioral sciences literature” (Rogosa et al. 1982, p. 726). See Möbus & Nagl, 1983, who suggest an integrated statistical model.

5. Although this affect-definition of attitudes is fundamental and accepted by most scientists, in many attitude scales we find items asking the respondent to give descriptions rather than evaluations (Cloetta, 1975, p. 43; Dann et al., 1978, pp. 368ff.). The importance of a clear distinction of Is and Ought statements in attitude research has been frequently underlined (Fishbein & Ajzen, 1975, p. 12; Triandis, 1975, pp. 4ff.; Deutscher, 1973, p. 316). Dippelhofer-Stiem (1983) has shown that this distinction is indeed of great empirical relevance when assessing the impact of the university environment on students’ development.

6. For example, the committee of the American Council on Education in the Fifties explicitly stated that it was “interested only in whether or not individuals accepted or rejected items” (Dressel & Mayhew, 1971, p. 218). See also Cloetta, 1975, pp. 38ff.; Müller-Fohrbrodt & Cloetta, 1975, p. 200; Dann et al., 1978, pp. 87ff.; Scott, 1968; Meinefeld, 1977, p. 190; Lippert & Wakenhut, 1978, p. 87.

7. Similarly Wilson: “Attitudes must show variation between individuals and between cultures, i.e. they relate to issues upon which people disagree” (1975, p. 95). This dependency of attitude or personality measurement on group or population scores is one of the main shortcomings of classical testing theory, yet most modern testing theories, have no remedy for it (Wakenhut, 1974).

8. Chickering, 1981a, p. 2. For further studies on higher education in the tradition of Kohlberg and Perry, see Haan et al., 1968; Collier et al., 1974; Marton & Säljö, 1976; Parker 1978; Rest, 1979a; Huber, 1980; Whiteley, 1980; Gilligan 1981.

9. Rest (1979b, part 3, pp. 10, 18). This data refers to the P-score which is
derived from answers to the *Defining Issues Test* (DIT). It is obtained by calculating the relative frequency of postconventional judgments (see Rest, 1979a).

10. For an interpretation of the profiles see Lind & Wakenhut (in this volume).

The research reported herein has been supported by the Deutsche Forschungsgemeinschaft as part of the ongoing longitudinal research project “University Socialization” at the University of Konstanz (see ch. 3, n. 4). I like to thank Hanns-Dietrich Dann, Roland Wakenhut, and Tom Wren for their comments on earlier versions of this chapter.
5. Moral judgment Competence and Political Learning

Horst Heidbrink

How political are moral dilemmas?

A fascinating question: If Lawrence Kohlberg's stage theory of the development of moral judgment were applied to political judgments, could politically controversial questions be morally decided in an "objective" manner? To establish the underlying moral stage, one would only need to analyze the arguments actually given – e.g., for and against squatting, nuclear power plants, deployment of medium-range missiles. The morally right side of the debate would then be the one whose proponents' arguments corresponded to the higher stage of moral development.

Such a procedure would have to assume that Kohlberg's stage theory applies to political judgments as well as to moral ones. However, difficulties arise at this point, since we lack a precise differentiation between moral and political judgments. For instance, Lockwood (1976, p. 317) limits his discussion to features which have the same impact on political questions as on moral ones. Kohlberg himself (1977, p.16) is obviously of the opinion that the development of political judgment – seen from a psychologist's point of view – represents a part of moral development. His view apparently is that political and moral education are largely identical. If common points and overlappings are emphasized in this way, there is, indeed, the danger that differences will be lost sight of.

The structural differences between real political problems and moral dilemmas, which are of crucial importance for Kohlberg's approach, are essential to the question of the transferability of the theory.

In Europe, a woman was near death from a rare form of cancer. There was one drug that the doctors thought might save her, a form of radium that a druggist in the same town had recently discovered. The druggist was charging $2,000, ten times what the drug cost him to make. The sick woman's husband, Heinz, went to everyone he knew to borrow the money, but he could get together only about half of what the drug cost. He told the druggist that his wife was dying and asked him to sell it cheaper or to let him pay later. But the druggist said, "No." So Heinz got desperate and broke into the man's store to steal the drug for his wife. (Kohlberg, 1976, p. 41)

The fictitious setting of the story (somewhere "in Europe") makes it pointless to ask questions about the social causes of the precarious position of the hus-
band – just as it would be pointless to propose specific political solutions, such as introducing legislation for health insurance. The Heinz story is a dilemma precisely because all alternatives are blocked. There is only one choice left to the husband: the death of his wife or theft.

Are moral dilemmas therefore of an apolitical nature, since they refer to fictitious situations which have little to do with reality? For purposes of comparison, let us consider a real problem situation, as reported in the German news magazine Der Spiegel (No. 23, 1981):

Every month for two and a half years the compulsory health insurance provider AOK in Euskirchen receives a reminder. The sender is the Institute for Experimental Hematology and Blood Transfusion Services of Bonn University, which is insisting on the refund of the treatment costs for a single patient. During the last thirty months, a debt of 23,949,263 DM has accumulated for the man. The health insurers from the Rhine Area keep refusing to pay.

So even the German health insurance system does not totally prevent the kind of problems contained in Heinz's dilemma. Contrary to the Heinz story, however, the situation in the case of the hemophiliac blood recipient remains ambiguous. The Spiegel article raises factual questions, whereas in the Heinz dilemma clear factual statements are made. As a consequence of its involvement in a barely understandable health system, the hemophiliac case shows high complexity and little transparency. To make it structurally comparable to the Heinz story, all unclear aspects would have to be clarified through additional information or additional assumptions, and the possible actions would have to be reduced to a set of clear-cut alternatives.

So the peculiarity of a moral dilemma lies less in its content than in its structure. Therefore, the fact that dilemmas with political contents can be constructed without special difficulties is in no way sufficient as a condition for the applicability of moral judgment theory to the political sphere. One of the most important structural particularities of moral dilemmas consists in the fact that they are in principle not optimally solvable: there is no possible action which does not produce some undesirable consequences. In the Heinz story, the husband has no chance to rescue his wife and also to satisfy the druggist. He has only two very imperfect options: to rob the pharmacist and so rescue his wife, or to refrain from stealing and thus let his wife die. Whatever alternative the subject chooses, the consequences are comparably unfortunate. For this reason Kohlberg ascribes little significance to the subject's actual choice. Pro- and con-decisions are of the same value at each stage, in that a dilemma is treated as though it were completely symmetrical. Information about someone’s stage is, therefore, to be taken from his reasons, not from his chosen course of action.
In practice, however, it is often rather difficult for the subject to understand a dilemma’s socio-moral structure, since this becomes obvious only in the course of a confrontation with arguments belonging to the next higher stage.

From the point of view of the next stage, pro and con reasons seem equally inadequate. From the point of view of one's own stage, the chosen option is, however, sufficiently legitimated. The cognitive conflict which may arise when one is confronted with arguments of the next higher stage would, therefore, conform to the insight that one has been led into the dilemma by one's own reasoning. For a person to face this conflict directly, and not try to avoid it by mobilizing defense or repression mechanisms (cf. Döbert & Nunner-Winkler, 1978) may indicate that he or she is already in transition from one stage to the next (cf. Kohlberg, 1969, p. 403). Presumably, the more determined and assertive one's arguments for or against an option in a symmetrical dilemma are, the more firmly one is bound to the stage one has attained. This “assertiveness of opinion” might be an important indicator for predicting the behavior that will manifest itself in the face of new information. It may be that persons showing little assertiveness are often in the middle of a reorientation process and hence more ready to accept new information. On the other hand, persons who usually judge in an extreme manner regard new information as a threat to whatever equilibrium has been achieved between their cognitive structure and their social environment.

This admittedly very speculative interpretation of the feature here called “assertiveness of opinion” refers to the symmetrical structure of moral dilemmas. However, political problems very seldom exhibit such symmetry. Consequently, in the absence of special information, we cannot assume that the decision and the stage of reasoning are independent of each other. A further essential difference has already been mentioned: the structure of a moral dilemma is set up, by the person making the judgment, unlike that of political problems, in which, to a large extent, the specific situations are formulated in advance, or “preconstructed.” This preconstruction can determine the stage up to which a problem is symmetrical, or the stage from which an asymmetry is found, such that reasons for certain decisions can no longer be given. Therefore, it is difficult to assess political arguments in an “objective” manner. Since a classification according to Kohlberg’s stage system is neutral regarding the validity of the assumptions underlying the given arguments, the quality of political judgments does not depend only on their moral stage, but also, and equally, on its congruence with reality (cf. Schneider, 1982).

Political learning constitutes an interactive process between the individual and his environment. When regarding identical aspects of reality, individuals may develop quite different internal representations. Recognition is determined not only by the objective structure of the “outside world” but also by the individually formed cognitive structure, i.e., the “inside world.” For interaction
with the political or social environment, the individual's stage of moral judgment is especially important. It may be assumed that, according to the level of moral development, political problem situations can be interpreted in various ways. New information is understood in a manner which allows it to be easily assimilated at the level of insight appropriate to one's stage of moral reasoning. The purpose of this procedure is to keep intact the balance between the individual and his environment once that balance is reached. We are more inclined to regard the environment in terms of our former experience than to keep changing ourselves to keep pace with the changing demands of the environment. The accommodation processes which are necessary for the further development of the moral judgment structure affect one's own self-interpretation and, hence, are risky (cf. Heidbrink, 1979, p. 18). The more strongly one is bound to an already-attained stage of moral judgment, the less able one will be to alter the quality of his or her political awareness. If this thesis holds, the ability for political learning closely depends on the process of moral judgment development.

Cognitive Complexity

Formally, the transformation of real political problems into moral dilemmas constitutes a reduction of complexity. By means of additional hypothetical assumptions, complex political circumstances can be brought to a "moral crux." At the same time, this transformation evens out inter-individual differences in the subjective judgment scopes. In the extreme case, political circumstances can be thought of in complete abstraction from the network of concrete social relationships, thereby eliminating all opportunities to interpret situations in a subjective manner. Problems are depoliticized, but at the cost of letting all the morally "nonproblematic" options fall victim to complexity reduction.

From a psychologist's point of view there are, according to Dörner (1976, p. 18), three main kinds of complexity reduction that are possible when solving problems: (1) abstraction, (2) formation of complexes, and (3) reduction. By means of abstraction, unessential variables are eliminated from the problem area, which result in a restriction of focus to those properties which are essential for the solution. Formation of complexes is, according to Dörner, the grouping of single properties into an integral whole. Reduction is carried out by considering a common basic property, which has as its consequences or symptoms other single properties, instead of considering all the single properties as a set.

When dealing with political circumstances, the appropriateness of the performed complexity reductions is often a matter of controversy. If, for instance, you do not distinguish between different reactor types when discussing nuclear
power plants, one person may accept the omission as a reasonable abstraction, but another may seize on it as a generalization that negates your argument. The formation of complexes can be even more explosive when applied to political circumstances, for it determines the consideration’s “degree of resolution.” The stronger the formation of complexes, the lower the degree of resolution. A very high degree of resolution leads to a “microscopic” view of the problem, whereas a low degree of resolution creates distance and allows the overall design to come into view. Different kinds of complex forming may lead to problem definitions which may have almost nothing to do with each other. Then the problem solutions will look correspondingly different. In the nuclear power debate, some demand a “general change of view” and trust in alternative forms of energy, economy, and lifestyle, while others counter that security systems in nuclear power plants are becoming more and more sophisticated. The most rigorous, and at the same time politically most risky, of these three ways of reducing complexity is the one which Dörner simply calls “reduction.” In the extreme case, all political problems can be reduced to a single common “basic evil.” As an example, Dörner mentions the reduction of all economic and political difficulties in the twenties and thirties to a “conspiracy of Zionism.”

In politics, reductions of complexity may be seen only partially as individual processes, because given reductions are often adopted unreflectively. Individual persons seldom have either the opportunity or the ability to see political circumstances in their full complexity. For example, whatever is presented in the media as “politics” is necessarily more or less reduced in advance. If somebody is unable to recognize such reductions for what they are, he is helpless and has to take them at face value.

The ability to deal with complex information in a complex manner is treated in the theory of cognitive complexity (Harvey et al., 1961; Schroder et al., 1975; Mandl & Huber, 1978; Streufert & Streufert, 1978). Huber (1982) gives a comprehensive treatment of the importance of this approach for political learning processes. As in the theory of moral judgment, learning is seen as an interaction between the individual and his environment. The actual information-processing capacity results from interactions between the individual cognitive structure and certain properties of the situation. In the present context, the most important property is the complexity of information, for example, the complexity of a political message. In empirical investigations, an optimal learning environment has proved to be one which is moderately complex with respect to the degree of cognitive structuredness. This relationship corresponds to a U-curve: either too much or too little complexity in the environment leads to a decrease in the actual information-processing capacity.

In contrast to real political problems, moral dilemmas show little information complexity and therefore require only a relatively modest processing capa-
city. Of course, they can (and should) evoke very differentiated moral considerations, but these are processes which should be distinguished from direct information processing. If students are confronted with moral dilemmas in a politics course, their moral judgment can be improved (Blatt & Kohlberg, 1975), but their ability to cope with complex political circumstances remains largely unaffected.

According to our theoretical opinion, three “levels of understanding,” marked by specific difficulties or learning impediments, can be analytically distinguished in political learning processes. The first level pertains to the form of presenting political circumstances, i.e., their surface structure. On this level, difficulties of understanding may be avoided through the choice of a different form of presentation which is more suitable for the learner. The concept of understandability developed by Langer et al. (1974) refers to this level. Learning impediments on the second or third level are more serious, because they are caused by the detail structure and thus do not depend on the way the facts are presented. On the second level, the understanding does not fail because of an unnecessarily complicated description, but because of the actual complexity of the facts. Merely simplifying the form of presentation will not be sufficient here; the complexity of the facts themselves must be reduced. The third level pertains mainly to the arguments given for possible problem solutions or the justifications of the consequences of actions, and is therefore associated with the stage of moral judgment. If the moral stage of the learner lies more than one stage below the stage of reasoning presented, an adequate understanding is impossible (Kohlberg & Turiel, 1971, p. 450).

With respect to the learner, these three levels can be considered a chain of barriers or filters. The first filter reduces descriptions which are too complicated, the second reduces factual complexity which is too high, and the third reduces moral arguments which are too pretentious. We have tried to show that moral dilemmas are so constructed that they usually pass smoothly through the first and second filters. Indeed, they require as direct a passage as possible to the third filter, i.e., to the person's moral judgment stage. With political problems this does not happen, because the first two filters are sensitive to real political circumstances producing partial or even complete blockage well before moral judgment is involved. It should now be clear why political education cannot be reduced to moral education. Moral judgment is a necessary but insufficient condition for the development of political judgment. It is important that students experience complexity as a constitutive property of political relations and of political thinking, and that they learn to cope properly with this complexity.
Moral judgment, cognitive complexity and political learning

The present investigation has been performed as a pilot study preceding a more extensive project dealing with questions of the use of media in political classes at vocational schools. It was designed to test instruments for measuring the stage of moral judgment and the degree of cognitive structuredness, and to assess the importance of both properties for political learning processes. Both properties are treated as independent variables, while the learning performance and the learning progress serve as dependent variables. Therefore the goal of this investigation was not to verify instructional possibilities for promoting cognitive and moral development, but to prove that both dimensions of development influence the gathering and processing of political information.

The investigation was performed in the winter of 1979 in 12 classes of vocational schools with a total of 294 students. In order to provide a comparable learning situation for all students, an educational television film from the series Politik Aktuell (Current Politics) of the German broadcasting company WDR was used. The film is entitled Organe aus zweiter Hand (Secondhand Organs) and discusses a controversial legislative proposal concerning organ donations. Lind's Moralisches Urteil Test (MUT; see Lind, 1978a; Lind & Wakenhut, in this volume) was employed to determine the moral judgment competence of the students. Two stories from the MUT were used, that of the workers and that of the physician. The degree of cognitive structuredness was determined through two incomplete sentences taken from The Paragraph Completion Test (PTC; Schroder et al., 1975), namely, “When I am criticized [. . .]” and “When I am in doubt [. . .]” The teaching goals of the educational television film were pinpointed by means of a related analysis (cf. Heidbrink, 1978). A teaching-goal-oriented test, LOT, based on these teaching goals, was devised specifically for this study to measure prior knowledge and learning progress. The test consisted of 15 multiple choice problems, which were given as a pretest to the experimental group and as a posttest to all students. In order to examine a possible sensitizing influence, the control group's pretest did not contain any questions related to the film.

Findings on Moral Judgment Competence

Further details of the investigation design and other procedures will not be discussed, since this would be of no direct relevance to the results described here. However, a more detailed look at the findings from the study will allow us to compare our results with those of the other studies described elsewhere in the present volume. Table 1 shows median and standard deviations for the stage preferences. For our purposes all four stage-specific items of the stories of the
workers and the physician were averaged.\textsuperscript{4} As expected, the medians depend on the theoretical order of stages. A clear increase in the values can be noticed as subjects advance to the next moral level, that is, at Stage 3 (conventional level) and at Stage 5 (postconventional level). For the analysis of the cognitive aspect of moral development, two measures were computed. At first, the degree of moral differentiation (the so-called “Stage factor”) was determined by Lind's (1978a, 1985d) method of the intra-individual decomposition of variance components (see also Lind & Wakenhut, in this volume). Additionally, intra-individual rank correlations (Kendall's tau) between the stage preferences and the theoretical order of stages were computed (Heidbrink, 1982, p. 86).

<table>
<thead>
<tr>
<th>Stage</th>
<th>Medians</th>
<th>Standard deviations</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1</td>
<td>1.6</td>
<td>0.74</td>
<td>273</td>
</tr>
<tr>
<td>Stage 2</td>
<td>1.5</td>
<td>0.75</td>
<td>274</td>
</tr>
<tr>
<td>Stage 3</td>
<td>2.0</td>
<td>0.72</td>
<td>269</td>
</tr>
<tr>
<td>Stage 4</td>
<td>2.1</td>
<td>0.63</td>
<td>276</td>
</tr>
<tr>
<td>Stage 5</td>
<td>2.6</td>
<td>0.63</td>
<td>271</td>
</tr>
<tr>
<td>Stage 6</td>
<td>2.6</td>
<td>0.64</td>
<td>265</td>
</tr>
</tbody>
</table>

\textbf{Table 1} Medians and Standard Deviations of Stage Preferences (MUT).

The analysis of variance shows that the Stage factor explains on average the highest proportion of the response variance, whereas the Story factor, i.e., differential acceptability due to the dilemma context, explains the lowest. Hence, when assigning preference values, the students orient themselves more according to the stage-specific quality of the given reasons than to the congruity of the reasons with the direction of their decision (the Pro-Con factor). Systematic differences in the responses to both stories seldom occur.
Despite the nearly identical numerical values for the Stage factor and the rank correlation tau, it should be emphasized that two clearly different operational realizations of moral judgment are involved. The Stage factor constitutes an intra-individual measure of variance which responds in the same manner to each kind of consistent distinction among arguments assigned to different stages. Because six stages allow $7^6 = 720$ different sequences, the maximum value 1 may be achieved theoretically by the same number of sequences. The rank correlation, instead, yields the optimal value $+1.0$ only if the empirical rank order of preference values is exactly equal to the theoretical order of stages. The correlation of these two measures was $r (\text{tau}) = .73$ ($N = 256$). The correlation value allows the conclusion that the variance, which is explained by the Stage factor, may indeed reflect Kohlberg's stage hierarchy. A model data analysis, which included the calculation of tau-values for each possible sequence of stages, also revealed a very good congruity of the empirical response structures to a theoretical ideal structure (cf. Heidbrink, 1984). After all, these findings may be considered a strong substantiation of both the validity of the MUT and the correctness of the moral judgment theory. This also follows from the close resemblance of our results to those of others using the MUT (cf. Lind, 1985d).

Correlations between All Three Variables

Before examining the influence of moral judgment and cognitive structuredness on the learning processes initiated by the television film, it was necessary to determine whether the students had made any measurable learning progress. Indeed, in the experimental group a highly significant difference between pretest and posttest values of the learning-goal-oriented test was obtained. A comparison with the control-group’s posttest values revealed that the
sensitizing influence of the pretesting could be statistically identified, but only within an order of magnitude which was negligible compared to that of the learning progress caused by the television film alone. Table 3 contains the correlations of the MUT and the PCT with the learning-goal-oriented learning achievements.5

<table>
<thead>
<tr>
<th>Learning Achievement</th>
<th>Pretest</th>
<th>Posttest</th>
<th>Progress (diff.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MUT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage factor</td>
<td>0.25**</td>
<td>0.32**</td>
<td>0.22**</td>
</tr>
<tr>
<td>Rank correlation tau</td>
<td>0.29**</td>
<td>0.41**</td>
<td>0.28**</td>
</tr>
<tr>
<td>Extremity of opinion</td>
<td>– 0.03</td>
<td>– 0.29**</td>
<td>– 0.31**</td>
</tr>
<tr>
<td>PCT</td>
<td>0.41**</td>
<td>0.45**</td>
<td>0.26**</td>
</tr>
</tbody>
</table>

**p = .01; n = 140

Table 3 Correlations between MUT, PCT, and Learning Achievement.

The correlations in Table 3 fully support our hypotheses. Both moral judgment measures show a highly significant correlation with all three learning-goal-oriented achievement measures. As expected, the students' degree of cognitive structuredness also correlates positively with learning achievements. Remarkably, the empirical results are compatible with the rather speculative thesis that extremity of opinion in the MUT can be considered an indicator of one's fixation at a certain stage and, therefore, as a measure of one's openness to new information. The insignificant correlation with the pretest achievements may be explained as a consequence of the time lag between the pretested learning processes, which of course must have been performed before the investigation started, and the determination of the extremity. This supports the opinion that extremity is a property which depends on the actual level of development.7

Pedagogical Implications

The results of our study suggest that, corresponding to our theoretical considerations, cognitive complexity and moral judgment constitute important preconditions of political learning processes. Although the empirical evidence covers
only a small section of the theoretically postulated relationships, it nevertheless suggests the notion that the chosen approach opens a promising path which may lead to a better understanding of political learning processes.

Pedagogically, it is important to take into consideration the students' actual cognitive and moral development stages. Of course this does not mean that teachers need to determine each student's development level by means of extravagant testing procedures before teaching about political issues. In our opinion, "indirect" methods of taking the students' learning readiness into account seem less costly and at the same time more appropriate to the goals of political education. Teachers need lesson plans and other teaching materials designed to offset whatever difficulties might arise from the students' developmental differences.

Such lessons could be based upon moral dilemmas which, because of their above-mentioned lack of complexity, involve special cognitive preconditions. But the lessons must not be restricted to fictitious situations nor to the level of "private" morals. Hence, moral dilemmas must eventually be " politicized." This should be done by the students themselves in the course of a discursive process. In small heterogeneous groups it might be equally important to stimulate moral judgment through arguments from the next higher stage, and to create complexity by discussing the conditions which have been introduced in the situation of the story. This type of procedure closely resembles that of political games, although the goal of describing reality as well as possible should never be forgotten.8

Taking complexity into account offers the opportunity to utilize Kohlberg's approach for the purpose of political education without falling victim to the danger of "depoliticization." The more the student realizes that complexity is a constitutive property of politics, the less he or she risks mistaking reductions for reality itself. Making this awareness vivid is especially important for the development of political judgment, because otherwise reductions could lead to a rigid conception of the world, a conception that is politically dangerous as well as naive and undifferentiated. In contrast, the knowledge that unsubstantiated assumptions about reality may lie beneath one's view of political problems seems a promising precondition for openness to other lines of argument.
Notes

1. For the measure “extremity of opinion” (Bilanzextremität) in the MUT, see Lind (1978a, p. 196).
2. We refer to the project Politische Bildung und Schulfernsehen in der Berufsschule (political education and educational television in vocational schools), carried out at the FEOIL-Institut für Medienverbund und Mediendidaktik by W. Hagemann, M. Schneider, and the author.
3. For a comprehensive description of the results, see Heidbrink (1982).
4. For vocational students the response scale was divided into five options instead of the original nine. For the evaluative calculations the raw values were transformed into values ranging from 0 to 4. For the sake of better comparability, the median in each case is given instead of the arithmetic mean. The two measures differ only slightly here.
5. The given correlations refer to the experimental group. Due to missing data, some correlations are based on fewer cases (minimum: 110).
6. The learning progress was determined by means of the regression differences.
7. For a description of these multivariate relationships, see Heidbrink (1982, p. 101).
8. On the basis of the procedure which is only roughly described here, a curriculum covering human rights education was realized in cooperation with the German broadcasting company WDR. These activities were financially supported by the Ministry of Culture of the German federal state of Nordrhein-Westfalen.
References


---

3 Editor’s note: I have not tried to tease out the references that are particular for the articles re-edited here. Also more recent publications related to the topic of these papers have been added to the list of references.
References


References


References

References

References


References


References


References


References


Lind, G., 1987, Kohlberg auf dem Prüfstand – Ein fiktives Gespräch über Schule, Demokratie und kognitiv-moralische Entwicklung. (Benchmarking Kohlberg – A fictitious conversation on schools, democracy and cog-
References


Lind, G., 2000. O significado e medida da competencia moral: um modelo do duplo aspecto da competencia moral. (The meaning and measurement of...
moral judgment competence revisited – A dual-aspect model. [Psicologia Reflexão e Critica, 13(3), in press. (ISSN 0102-7972).]
References

References


References

References


147


Appendix: Authors of the original edition

Tino Bargel, Research Sociologist, Forschungsgruppe Hochschule of the Social Science Faculty, University of Konstanz. Major fields of research are education and socialization, social ecology and regional development, politics and social change. Co-editor and co-author of "Bildungschancen und Umwelt" ([Educational opportunities and environment] 1973, published by Westermann), "Gesamtschule und Gemeinde" ([Comprehensive school and community] 1979, Schöningh), and "Studien situation und studentische Orientierungen" ([Conditions of study and students' orientations] 1984, Bock). Address: FB Sociology and History, University of Konstanz, 78457 Konstanz, Germany.


Gerhard Bohm, Dr. Clinical Psychologist. Working with handicapped persons and disturbed families. Fields of research are political participation and moral reasoning.

Rainer Döbert, Dr., PD., Sociologist. Major fields of research are identity formation during adolescence, development of moral consciousness, and religious evolution. Author of "Systemtheorie und die Entwicklung religiöser Deutungssysteme" ([System theory and the development of religious systems of interpretation] 1973, Suhrkamp), and "The role of stage models within a theory of social evolution" (in R. Harré & J. Jensen, eds., The philosophy of evolution, 1981, The Harvest Press), and co-author of "Formal and material role taking" (1985, Human Development).

Hans A. Hartmann, Dr., Professor of Psychology at the Department of Economic and Social Sciences, University of Augsburg. Major fields are Economic Psychology, Social Psychology and Psychodiagnostics. Special interests in ethics and its application to psychology. Author of "Psychologische Diagnostik" ([Psychological diagnostics] 1973, Kohlhammer), and Co-editor of "Lehrbuch der Holtzman Inkblot Technique," 5 volumes ([Textbook of the Holtzman Inkblot Technique] 1977, Huber), "Advances in Economic Psychology" (1981, Meyn), and "Psychologische Begutachtung" ([Psychological
Horst Heidbrink, Psychologist. Dr., Lecturer at the "Fernuniversität" (open university, university with regular admission requirements but no classrooms; mailed material instead) in Hagen. Formerly research assistant at the FEoLL-Institute, Paderborn. Majors field is Educational Psychology. Special interests in the psychology of social interaction, cognitive-moral development and Educational Psychology. Co-editor and of author of "Kognition und Moralität" ([Cognition and morality] 1982, Leske), and author of essays on "Analyse von Lernzielen und Lernstrukturen als Evaluationsverfahren" ([Analysis of aims and structures of learning], in K. Benecke et al., eds., Schulfernsehen in Theorie und Praxis, 1981, Leske) and "Zur Validität der moralischen Urteilstheorie von L. Kohlberg" ([On the validity of Kohlberg's moral judgment theory], in K. Ingenkamp, ed., Sozial-emotionales Verhalten in Lehr- und Lernsituationen, 1984, EWH Rheinland-Pfalz, Landau).

Thomas Krämer-Badoni, Dr., Professor of Sociology at the University of Bremen. Formerly assistant researcher of the Max-Planck-Institute at Starnberg. Major fields of research urban and regional planning. Author of "Zur Legitimität der bürgerlichen Gesellschaft" ([In the legitimacy of the bourgeois society] 1978, Campus) and co-editor of "Die Kommune in der Staatsorganisation" ([The community within the state organization], 1974, Suhrkamp).

Appendix


Gerhard Portele, Dr., Professor of Higher Education. Mains fields are Motivation and Learning, Moral and Clinical Psychology. Author of "Widerspruchsfreiheit und Vollständigkeit als Eigenschaften kognitiver Systeme" ([Absence of contradiction and completeness as properties of a cognitive system], in H. Uekert & D. Rhenius, eds., Komplexe menschliche Informationsverarbeitung, 1979, Huber), and editor of "Sozialisation und Moral" ([Socialization and morality] 1978, Beltz). Address: Interdisziplinäres Zentrum für Hochschuldidaktik, Universität Hamburg, Sedanstr. 19, D-2000 Hamburg.


Michael Schenk, Dr. habil., Professor at the Institut für Publizistik der Universität Mainz. Majors field is communication research; special interests in moral judgment and political participation. Author of "Publikums- und Wirkungsforschung" ([Research on effects of mass media] 1978, Mohr), "Kommunikationsstrukturen in Bürgerinitiativen" ([Communication structures in single issue movements] 1982, Mohr), and "Soziale Netzwerke und Kommunikation" ([Social networks and communication] 1984, Mohr).

André Schläfli, Dr., Psychologist in the Neurological Clinic of the University of Zürich. Major fields are moral education, neuropsychology and the psychology of emotions. Author of "Förderung des sozial-moralischen Kompetenz: Eine Interventionsstudie mit Lehrlingen" ([Fostering of social-moral competence: A intervention study with trainees] 1984, unpubl. doctoral dissertation, University of Fribourg, Switzerland); co-author of "Does moral education improve moral judgment? A meta-analysis of intervention studies using the Defining Issues Test" (1983).

Roland Wakenhut, Dr., Professor of Applied Psychology. Major fields are Political Psychology, Military Psychology, and Statistics. Author of "Messung gesellschaftlich-politischer Einstellungen" ([The measurement of socio-political attitudes] 1974, Huber), and "Über die Einbeziehung von Situationen in psychologischen Messungen" ([On the role of situations in
psychological measurement] 1978, Lang), and co-editor of "Handwörterbuch
der Politischen Psychologie" ([Dictionary of Political Psychology] 1983,
Westdeutscher Verlag) and "Selektion oder Sozialisation" ([Selection or so-
Contents of the original version, 1985

Preface .................................................. IX

Foreword by Lawrence Kohlberg ............................. XV

Theory and Methods

Philosophical Notes on Moral judgment Theory
by Jürgen Habermas ........................................ 3

The Theory of Moral-Cognitive Development:
A Socio-Psychological Assessment
by Georg Lind ............................................. 21

Moral Competence and Democratic Personality
by Georg Lind, Johann-Ulrich Sandberger,
and Tino Bargel ........................................... 55

Testing for Moral judgment Competence
by Georg Lind and Roland Wakenhut ....................... 79

Moral judgment and Its Social Milieu

Social Inequality and Moral Development
by Hans Bertram ............................................ 109

Value Change and Morality by Rainer Döbert
and Gertrud Nunner-Winkler ............................. 125

The Thin Line Phenomenon: Helping Bank Trainees
Form a Social and Moral Identity by Fritz Oser
and André Schläfli ........................................... 155

Attitude Change or Cognitive-Moral Development?
by Georg Lind .............................................. 173

How Scientists Think about Science and Morality
by Gerhard Portele ........................................... 193

155
Morality and the Military Life-World
by Thomas Krämer-Badoni and Roland Wakenhut .......... 205

Segmentation of Soldiers' Moral judgment
by Rainer Senger ................................................ 221

Single-Issue Movements: Political Commitment and
Moral judgment
by Michael Schenk and Gerhard Bohm ...................... 243

Moral Judgment Competence and Political Learning
by Horst Heidbrink ............................................. 259

Philosophical Epilogue: Phylogeny and Ontogeny

What Is Social About Morality? Morals, Morality, and
Ethics in Social Science Perspectives
by Hans A. Hartmann ........................................... 275

Contributors ........................................................ 296

References .......................................................... 299

Index of Names .................................................. 319

Index of Subjects ................................................ 325