



Konstanz, 1980

Korrigierte und ergänzte Neuauflage 2002, 2019

Hinweis des Autors zum Neudruck 2019: Neue Veröffentlichungen zu diesem Themen finden sich im Internet: <http://www.uni-konstanz.de/ag-moral/>

Eine Zusammenfassung der neuesten Forschung und der daraus inzwischen resultierenden Anwendungen auf dem Gebiet der Moral- und Demokratie-Erziehung findet sich in:

Lind, G. (2019). Moral ist lehrbar. Wie man moralisch-demokratische Kompetenz fördern und Gewalt und Betrug reduzieren kann. Berlin: Logos.

Universität Konstanz
Zentrum I Bildungsforschung
Sonderforschungsbereich 23

Georg Lind

ENTWICKLUNG DER STUDENTISCHEN PERSÖNLICHKEIT:

BEITRÄGE ZUR THEORIE UND METHODE DER
SOZIALISATIONSFORSCHUNG

Arbeitsunterlage 67

1980

Forschungsgruppe Hochschulsozialisation:

Peisert, Hansgert, Professor, Dr.rer.pol.
(Projektleiter)

Bargel, Tino Ernst, Dipl. Soz.
(stellv. Projektleiter)

Dippelhofer-Stiem, Barbara, Dr. rer. soc.

Framhein, Gerhild, M.A.

Lind, Georg, Dipl.Psych.

Sandberger, Johann-Ulrich, M.A.

Walter, Hans-Gerhard, M.A.

Die Forschungsgruppe *Hochschulsozialisation* befaßt sich mit Entwicklungen und Problemen im Hochschulbereich.

Im Mittelpunkt steht das Projekt "Bildungsbiographien und Daseinsvorstellungen von Akademikern" (*Hochschulsozialisation*). Es untersucht im Rahmen einer international vergleichenden Längsschnittstudie die Voraussetzungen, Abläufe und längerfristigen Folgen der Sozialisation während des Studiums.

Das Projekt wird im Rahmen des Sonderforschungsbereichs 23 durch die Deutsche Forschungsgemeinschaft gefördert.

In Ergänzung zu diesem Projekt werden aktuelle Untersuchungen im Auftrag von Bund, Ländern und der eigenen Universität durchgeführt (z.B. "Entwicklungsprobleme neuer Universitäten", "Informationsverhalten von Studenten").

Eine zusammenfassende Darstellung der Forschungstätigkeiten und Veröffentlichungen der Forschungsgruppe finden sich im Wissenschaftlichen Arbeits- und Ergebnisbericht 1976-1978, Konstanz 1979.

INHALT [Seitenzahlen wie im Original]

Kapitel 1 :

Die zwei Ansätze zur Erfassung der Identität und ihre Integration -
Überlegungen zum Dilemma der Hochschulsozialisationsforschung 7 - 19

Kapitel 2:

Kognitiv-strukturelle Analysen von Einstellungssystemen 20 - 29

Kapitel 3:

Reliability of Tests or Structure of Personality? Empirical Findings on the
Problem of Interpreting Judgment Consistency 30 - 35

Kapitel 4:

Experimental Questionnaires: A New Approach to Personality Research 36 - 50

Literatur 51 - 56

Veröffentlichungsnachweis 58

Vorbemerkungen

In den programmatischen Aufsätzen über Persönlichkeit und Identität in der (Hochschul-) Sozialisationsforschung wird vielfach auf Probleme ihrer Grundlagen hingewiesen. Dies scheint jedoch kein spezifisches Problem dieses Forschungsbereichs zu sein, sondern Teil einer grundlegenden Problematik der Persönlichkeitsforschung und der Entwicklungspsychologie, wie sie gelegentlich bereits in früheren Jahren angesprochen wurde (K. Bühler, G. Allport, J. Loevinger, Cicourel u.a. m.). Hier wie dort ist ein wichtiges, immer wiederkehrendes Problem die Passung, beziehungsweise der Mangel an Passung von Theorie und Methode. Eine globale Zurückweisung der 'quantitativ' verfahrenen Sozialisationsforschung und der ihr verbundenen Konzepte, wie von manchen gefordert, läßt sich damit sicher nicht rechtfertigen; es mehren sich die Anzeichen, dass die unhinterfragte Akzeptierung von psychometrischen Vorschriften zu Paradoxien (z.B. dem Reliabilitäts-Veränderungs-Paradoxon) und zu einer Ausblendung der gesamten kognitiv-strukturellen Komponente der Persönlichkeitsentwicklung aus der Forschung führt. Am Beispiel der kognitiven Entwicklungstheorie des moralischen Bewusstseins ist diese Problematik in den letzten Jahren in ganzer Schärfe hervorgetreten.

Verfechter des Dogmas der 'klassischen' Psychometrie werfen diesem Forschungsparadigma die Verletzung psychometrischer Gütekriterien vor (u.a. Kurtines & Greif 1974; Phillips & Nicolayev 1978; zur Entgegnung vgl. Broughton 1978, Oser 1980, Lind 1980a). Diese Vorwürfe können jedoch nur dann als legitim gelten, wenn die ihnen zugrundeliegende, zumeist aber implizit bleibende Psychologie empirisch valide ist.

In den folgenden Aufsätzen, die alle 1980 zu unterschiedlichen Anlässen als Vorträge entstanden sind, wird der Versuch einer Integration der psychometrischen Position und der Kritik an ihr unternommen. In dem ersten Beitrag werden das der Psychometrie zugrundeliegende 'common trait' Modell und das immer wieder als Alternative angeführte 'internal structure' Modell diskutiert. Ich vertrete dort die These, dass die Kontroverse zwischen diesen, sich bislang unvermittelt gegenüberstehenden Modellen nicht durch den Austausch von Begriffen zu überwinden ist, dass vielmehr ein neues Verständnis von der Natur der Persönlichkeit gewonnen werden muß, durch das die Unzulänglichkeiten beider Ansätze lokalisiert werden können.

Der zweite Beitrag knüpft an kontroverse Themen der politischen Sozialisationsforschung an und weist (mit Beispielen aus der Forschung des Projekts Hochschulsozialisation) einige Implikationen des kognitiv-strukturellen Ansatzes auf. Es wird gezeigt, dass sich das von Converse entdeckte Phänomen der 'non-attitudes' durch die Persönlichkeitstheorie von Piaget und Kohlberg auf eine breitere Theoriebasis stellen läßt. Der dritte Aufsatz analysiert das oben erwähnte Reliabilitätsproblem, das Ausgangspunkt für die Notwendigkeit neuer Ansätze in der Persönlichkeitsforschung ist. Es werden Belege dafür erbracht, dass die ausschließliche, von der Psychometrie vorausgesetzte Rückführung von Antwortinkonsistenz auf Eigenschaften des Instruments nicht haltbar ist.

Der letzte Beitrag schließlich stellt einen neuen Ansatz der psychologischen Forschung, den "Experimentellen Fragebogen" dar. Ich entwickle dort, teilweise anknüpfend an ältere Vorstellungen (W. Stern, R. Cattell, G. Allport, J. Loevinger), das Konzept des strukturell-dynamischen Persönlichkeitsmerkmals ("dynamic-structural trait") und schlage neue Wege der Konstruktion, Auswertung und Bewertung von psychologischen Meßinstrumenten vor. Alle

Beiträge lagen bislang nur als Vortragsmanuskript vor. Der interessierte Widerhall hat uns veranlaßt, sie in dieser Arbeitsunterlage zugänglich zu machen.

Drei der Manuskripte sind zur Publikation in Sammelbänden angenommen, werden jedoch erst im Laufe des nächsten Jahres, zudem an ganz verschiedenen Stellen, erscheinen (vgl. den

Nachweis auf der letzten Seite). Aus den unterschiedlichen Kontexten, in denen diese Vorträge entstanden sind, folgt eine gewisse, auch sprachliche Heterogenität, die wir, ebenso wie inhaltliche Überschneidungen, in dieser Zusammenstellung belassen haben. Nur die Literaturhinweise wurden am Ende zusammengefaßt.

Zum Neudruck 1982

Neben einigen kleineren Korrekturen wurde der Veröffentlichungsnachweis auf den neuen Stand gebracht. Für neuere Arbeiten zum gleichen Themenkreis sei auf Lind [(2019)] verwiesen.

Zur Neuauflage 2000

Aufgrund der nachhaltigen Relevanz der dargelegten Forschungsergebnisse angesichts bildungspolitischer Debatten liegt diese Arbeitsunterlage nun in elektronischer Form erneut vor. Der Inhalt wurde unverändert wiedergegeben. Das Manuskript hat lediglich ein paar formale Änderungen in der Gestaltung erfahren und wurde nochmals korrekturgelesen.

Korrekturen und Ergänzungen 2019 in [eckigen Klammern].

Kapitel 1

DIE ZWEI ANSÄTZE ZUR ERFASSUNG DER IDENTITÄT UND IHRE INTEGRATION ÜBERLEGUNGEN ZUM DILEMMA DER HOCHSCHULSOZIALISATIONSFORSCHUNG¹

Einleitung

Ich möchte mit meinem Beitrag einige grundsätzliche Überlegungen zu dem Dilemma der Sozialisationsforschung anstellen, und an einem Beispiel erläutern, wie wir das Problem der Übersetzung theoretischer Fragestellungen in adäquate, will heißen in *theoretisch valide* Erhebungs- und Auswertungsdesigns zu lösen versucht haben. Auf dieses Dilemma wurde im Bereich der Hochschulsozialisationsforschung bereits mehrfach hingewiesen, so von Huber (1975) und von Nitsch (1975) auf der Tagung 'Sozialisation in der Hochschule' in Bielefeld im Frühjahr 1975, so auch schon von Riesmann (1963) in seiner Analyse des 'Jacob-Reports' (1957). In der Vergangenheit ist man dem Problem der Koordination von Theorie und Methode nicht selten ausgewichen; einmal in eine heftige methodendogmatische Kontroverse um die von vielen ebenso hartnäckig weiterbenutzten wie von einigen strikt abgelehnten Methoden der 'klassischen' Psychometrie; zum anderen in eine Reduktion des Problems auf die Frage nach den rechten Begriffen. Ich glaube es besteht Einigkeit darüber, dass keine der beiden Auswege, weder der methodendogmatische noch der begriffsscholastische, letztlich an dem Problem der adäquaten Inbeziehungsetzung von Theorie und Methoden vorbeiführen.

Die folgenden Überlegungen gehen davon aus, dass sowohl die Frage der adäquaten Methode als auch die Frage der geeigneten Begrifflichkeit noch offen ist. Sie finden nicht nur auf den Terminus 'Identität' ihre Anwendung, sondern grundsätzlich auch auf ähnliche Begriffe wie Persönlichkeit, Habitus, Selbst, Perspektive, Einstellung und Position, insofern sie jene strukturdynamische Organisation des individuellen Verhaltens meinen, in der sich der Prozeß der Sozialisation spiegelt (siehe Tabelle 1). Diese Begriffe sind so sehr mit Bedeutung aufgeladen, dass eine präzise Abgrenzung gegenwärtig kaum möglich scheint. Nicht wenige Autoren verwenden sie zudem synonym.

¹ Beitrag für die Arbeitsgruppe Sozialisation in und durch die Hochschule, 20. Deutscher Soziologentag in Bremen, 1980.

Tabelle 1
Konzepte vom Gegenstand der Sozialisationsforschung

Persönlichkeit	R. König (1969, Lexikon der Soziologie) G. Hartfiel (1972, Wörterbuch der Soziologie) G. Allport (1961, Pattern and growth of personality)
Habitus	J. Habermas et al. (1961/1969, Student und Politik) P. Bourdieu und Passeron (1971, Die Illusion der Chancengleichheit) L. Huber & G. Portele (1980, Hochschulsozialisation)
Identität	E. Erikson (1966, Identität und Lebenszyklus) R. Döbert & G. Nunner-Winkler (1975, Adoleszenzkrise und Identitätsbildung)
Selbst	G.H. Mead (1934, Mind, self and society)
Perspektive	H. Becker et al. (1961, Boys in white)
Position	W. Perry (1970, Intellectual and ethical development in the college years)
Einstellung	L. Thurstone (1931, the measurement of social attitudes) B. Cloetta (1975, Einstellungsänderung durch die Hochschule) Th. Newcomb & K. Feldman (1969, The impact of college on students)
Status	J. Meyer (1977, The effect of education as an institution)

Die zwei Ansätze: Gegenüberstellung und Kritik

Nimmt man positivistische Positionen aus, die das Theoretische im Forschungsprozeß leugnen und deshalb von dem Forschungsobjekt auch eigentlich keinen Begriff haben, dann können die bisherigen Versuche der Inbeziehungsetzung von Theorie und Methode zwei Ansätzen zugeordnet werden. Auf der einen Seite der methodisch "sauberen" datenmäßig breiten, aber theoretisch unbefriedigenden quantitativen Sozialisationsforschung; auf der anderen Seite der theoretisch reiferen, aber methodisch "unsauberen", auf wenige Fallbeobachtungen beschränkten, phänomenologischen Sozialisationsforschung. Ich will die erste – etwas vereinfachend und mehr ihr reales Vorgehen und möglicherweise nicht ihr Selbstverständnis beschreibend – mit Gordon Allport (1961) als AußenEffekt-Ansatz bezeichnen. Die zweite will ich (ebenfalls vergrößernd und in Allports Terminologie) dem Internale-Struktur-Ansatz zurechnen. Beide Ansätze will ich im Folgenden kritisch miteinander konfrontieren. Im Anschluß daran sollen unsere Vorarbeiten zu einer Integration beider Ansätze skizziert und am Beispiel des methodischen Konzepts des Experimentellen Fragebogens über Erfahrungen damit berichtet werden. Zuletzt will ich untersuchen, welche Folgerungen sich hieraus für den epistemologischen Standpunkt der Sozialisationsforschung ergeben.

Der Außen-Effekt-Ansatz umfaßt vor allem die Forschungen, die sich der klassischen Test- und Fragebogentheorie bedienen. Zwar wird von jenen, die diese Methodologie positivistisch unreflektiert einsetzen, bestritten, dass sich damit eine bestimmte theoretische Position verbindet (vgl. u.a. Hogan et al. 1977, S. 255), bei genauerer Hinsicht läßt sich eine solche aber unschwer ausmachen. Der Außen-Effekt-Ansatz geht von gemeinsamen Eigenschaften ('common'- trait' Modell) der Persönlichkeit aus. Er setzt voraus, dass jede individuelle Identität eindeutig im Wertebereich einer äußerlichen a priori definierten Kategorie lokalisiert werden kann.

Die wichtigste Konsequenz dieses Modells ist, dass die oft erheblichen Inkonsistenzen bei der Klassifikation – man erhält durchschnittlich nur etwa 10 Prozent konsistenter Merkmalsvarianz (vgl. Mischel 1968) – auf die Unzuverlässigkeit (Unreliabilität) des Meßinstruments zurückgeführt und nicht mit den untersuchten Personen in Verbindung gebracht werden. Diese angeblich rein meßtechnische Erklärung enthält offensichtlich eine weitere Implikation. Sie impliziert, dass die Größe der Inkonsistenzen weder mit dem gemessenen Merkmal noch mit der gemessenen Person zusammenhängt. In dieses Modell paßt nicht die Differenzierung der Eigenschaften nach individueller Bedeutung, saliency oder “affektiver Selbstidentifikation” (Reichwein 1980, S. 5). Die Möglichkeit von ‘non-attitude’, ‘non-traits’ oder ‘Nicht-Identitäten’ wird im Rahmen dieses Modells gar nicht erst als Möglichkeit in Rechnung gestellt. Für solche individuellen Idiosynkrasien ist das Meßmodell nicht sensibel. In dieser strengen Fassung – andere, weniger klare Fassungen sind meist empirisch gehaltlos – ist der Außen-Effekt Ansatz empirisch nicht mehr generell haltbar (Lind 1980b). Die reale Verhaltensorganisation entspricht diesem Modell nur dort, wo sie durch soziale Institutionen (Schule, Kultur, Rechtsnormen) in Übereinstimmung gebracht wird (vgl. z.B. “Intelligenz”).

Der Internale-Struktur-Ansatz

Der Internale-Struktur-Ansatz geht davon aus, dass jede Person eine idiosynkratische Identität besitzt, die allein aufgrund individueller Deutungsschemata (phänomenologisch) zu verstehen ist. Dieser Ansatz versucht deshalb den Sozialisanden von “innen” zu verstehen und nicht, ihn von “außen” mit einheitlichen Beschreibungskategorien zu erklären. In der Hochschulsozialisationsforschung verbinden sich damit vor allem Namen wie Howard Becker (et al. 1961), William Perry (1970) und Ference Marton (1978) (vgl. auch Brennan & Franke-Wikberg 1980).

Dieser Ansatz stößt aber dort, wo er streng angewendet wird, auf zwei unlösbar erscheinende Probleme. Zum einen würde gerade die von diesem Ansatz unterstellte Einzigartigkeit der individuellen Identität die Einbeziehung zahlreicher Fälle erfordern, also einen quantitativen Ansatz. Welche Schlüsse lassen sich für die hochschulische Sozialisation aus zehn oder zwanzig idiosynkratischen Interviews ziehen? Dieses Paradoxon ist gewissermaßen symmetrisch insofern wenige Fälle dann ausreichen würden, wenn nicht dieses, sondern das 'common-trait' Modell empirisch gültig wäre. Zum anderen muß der Anspruch als problematisch gelten, Identitäten von "innen", also ohne Rückgriff auf äußere Beschreibungskategorien zu verstehen. Bei genauerem Hinsehen zeigt sich, dass auch phänomenologische Forschung keineswegs ohne a priori Kategorien auskommt, wenngleich sie diese vielfältiger gestaltet und eher den Gegebenheiten des Einzelfalls anzupassen bereit ist. Würde die Äußerlichkeit dieser Kategorien jedoch geleugnet, dann entzöge sich dieser Ansatz jeder rationalen Kritik. Der phänomenologische Ansatz wäre dann nur eine Spielform des Positivismus und wie dieser unakzeptabel.

Anforderungen an eine Methodologie und Versuch einer Integration der zwei Ansätze

Ich will hier die These vertreten, dass die Hochschulsozialisationsforschung – will sie ihre derzeitige Krise überwinden – auf der theoretischen Ebene zu einer Synthese des Außen-Effekt und des Internalen-Struktur Ansatzes und auf der methodischen Ebene zu einer Verbindung von kategorienanwendenden und kategorienauffindenden Verfahren gelangen muß. Diese Integration muß – im Gegensatz zu dem 'common-trait' Modell anerkennen, dass sich Personen nicht nur hinsichtlich des Ausmaßes unterscheiden, mit dem sie eine a priori definierte Eigenschaft besitzen, sondern auch danach, ob diese Eigenschaften für ihr Verhalten als Organisationsprinzip eine Bedeutung besitzen. Es gilt, zwischen der Ausprägung einer Eigenschaft und dem Grad ihrer Verhaltensrelevanz zu unterscheiden und damit den Konsistenzbegriff mit inhaltlicher Bedeutung zu füllen.

Im Gegensatz zum Internalen-Struktur Ansatz müssen wir aber auch anerkennen, dass wir als Forscher die Äußerlichkeit unserer Beschreibungskategorien nicht überwinden, sondern höchstens berücksichtigen können. In dieser Differenzierung ist die Wittgensteinsche Einsicht eingeschlossen, dass es ein absolutes Verstehen nicht geben kann, dass jedoch ein relativer Verständniszugewinn möglich ist. Voraussetzung für diesen Zugewinn ist, dass Theorie und Methodik so aufeinander abgestimmt sind, dass die Korrekturbedürftigkeit der Beschreibungskategorien erkennbar wird.

Dieser Anforderung konnten bislang beide Ansätze nicht hinreichend genügen. Entweder hat man, wie beim Außen-Effekt Ansatz, den Mangel an *kategorialer Überlappung* (Campbell 1963) nicht gesehen, weil er als rein technisch bedingt interpretiert wurde, oder, wie in beiden Ansätzen häufig der Fall, die individuelle Beschreibungskategorie als hypothetisch oder latent ins Reich des Irrealen verbannt.

Im Projekt *Hochschulsozialisation* haben wir ausreichende Erfahrungen mit den beiden Ansätzen, mit quantitativen wie qualitativen Verfahren, mit offenen Interviews wie mit psychometrischen Fragebogen. Wir müssen, wie jüngst Reichardt und Cook (1979) feststellen, dass sich viele der üblichen Gegenüberstellungen (vgl. Tabelle 2) dabei als irrelevant erwiesen haben. Es hat sich gezeigt, dass "qualitative" Methoden allein noch keine Qualität verbürgen und dass quantitative Methoden auf qualitativen aufbauen müssen. Es hat sich auch gezeigt, dass "offene Interviews" in der Durchführung und der Auswertung weit mehr der Standardisierung unterworfen werden müssen als das Adjektiv "offen" suggeriert und dass sich mit "geschlossenen Fragebogen" durch eine individuen-bezogene Konstruktion und Auswertung gegenüber der klassischen Psychometrie eine erhebliche Vergrößerung der kategorialen Überlappung, also ein besseres Verstehen erzielen läßt.

Wir haben darüber hinaus gesehen, dass hinsichtlich der Frage der Angemessenheit der Forschungsmethode innerhalb der Verfahren bei Interview und Fragebogen oft erheblich größere Unterschiede festzustellen sind als zwischen diesen Verfahren (vgl. hierzu auch Windolf 1979).

Tabelle 2
Die zwei Ansätze der Sozialisationsforschung*

1. Ansatz	2. Ansatz
Empirische Sozialforschung Surveyforschung	Phänomenologie Ethnographie
Modell der Identität:	
Außen-Effekt-Modell (common-trait)	Internale-Struktur Modell
Adjektivistische Unterscheidungsmerkmale:	
quantitativ	qualitativ
noumenal	phenomenal
objektiv	subjektiv
einfach	komplex
naturwissenschaftlich	geisteswissenschaftlich
deduktiv	induktiv
experimental	experiential
partikularistisch	holistisch
kausal	intentional
Typisches Forscherverhalten:	
erklären	verstehen
messen	beschreiben
beobachten	teilnehmen
Kategorien anwenden	Kategorien entdecken
generalisieren	explorieren

* In Anlehnung an Reichhardt und Cook (1979) und Marton (1978).

Auf dem Hintergrund dieser Erfahrungen schien es uns sinnvoll, eine Integration beider Ansätze zu versuchen, die ihre jeweiligen Nachteile vermeidet. Der neue methodische Ansatz sollte von einer Explikation des Identitäts- und Persönlichkeitskonzeptes ausgehen, damit dieses mit einer theoretisch validen Methode überprüft werden kann. Er sollte deshalb auch sensibel für Idiosynkrasien der realen Verhaltensmuster sein, um Unzulänglichkeiten der Theoriebildung erkennen zu können. Der hier bevorzugte dynamisch-strukturelle Identitätsbegriff impliziert, wie an anderer Stelle ausführlicher dargelegt (Lind 1980c), dass mit dem neuen Ansatz möglich ist, (a) erst für jede Person die Frage zu klären, ob und in welchem Maße ein vermutetes Persönlichkeitsmerkmal das beobachtete Antwortmuster zu einem Ganzen organisiert (Konsistenz), bevor (b) ihr eine Ausprägung dieses Merkmals zugeordnet wird (vgl. Tabelle 3).

Unterscheidungsmerkmal	Außen-Effekt-Ansatz	Internale Struktur-Ansatz	Integrationistischer Ansatz
Identität als Ausprägung auf äußeren Dimensionen	Ja	Nein	Ja, aber nur in bestimmten Fällen
Identität als Identität als idiosynkratische Organisation von Verhaltensweisen	Nein	Ja	Nein, aber für Idiosynkrasien sensibel
Erfassung - des dynamischen Aspekts der Identität	Ja	Nein	} Ja, aber beide als nicht trennbare, aber unterscheidbare Aspekte
- des strukturellen Aspekts der Identität	Nein	Ja	

Tab. 3: Unterscheidungsmerkmale der zwei Ansätze und ihrer Integration (siehe Text)

Das Konzept des "Experimentellen Fragebogens"

Wir haben mit dem Konzept des Experimentellen Fragebogens eine Forschungsstrategie entwickelt, die die vorgenannten Anforderungen erfüllt. Dieses Konzept bedient sich der Ökonomie und Transparenz der Fragebogentechnik und der erkenntnistheoretischen Vorzüge des experimentellen Designs. Gerade das Experiment wurde lange Zeit durch den Positivismus für Erwartungen vereinnahmt, die es ebensowenig wie jede andere Methode je erfüllen konnte. Entläßt man es aus diesem Kontext und stellt es statt dessen in den Dienst der Überprüfung von empirisch gehaltvollen theoretischen Aussagen, dann erweist sich das Experiment wie keine andere Methode als geeignet, komplexe Prozesse zu objektivieren und intersubjektiv nachprüfbar zu machen. Es entspricht in vollkommener Weise dem hermeneutischen Zirkel, der ausgehend von einzelnen Beobachtungen und der probeweisen Interpretation dieser Beobachtungen als Hinweis auf eine ihr Bedeutung verleihende Ganzheit in weiteren empirischen Beobachtungen Bestätigung und Widerlegung sucht, um sich, zwischen Beobachtung und Interpretation iterierend, der Isomorphie zwischen beiden zu vergewissern (vgl. Krausser 1968; Lind, Nielsen & Schmidt, 1976, S. 136).

Experimentelle Fragebogen bilden die Interaktion ab zwischen Person und Situation. Sie ermöglichen die Analyse eines Antwortmusters auf ein sorgfältig aufgebautes Fragemuster. Sie sind deshalb nicht mit üblichen Selbstbeschreibungsfragebogen zu vergleichen, sondern eher mit dem klinischen Interview, mit dem der Psychoanalytiker und der kognitive Entwicklungspsychologe dynamische Identitätsstrukturen in vivo studieren. Experimentelle Fragebogen sind gewiß weniger flexibel als diese, erlauben aber dem Befragten eine ähnlich differenzierte Antwortweise. Sie sind zudem nicht nur ökonomischer einzusetzen, sondern zeichnen sich auch durch größere Transparenz aus, was sowohl für ihre Handhabung als auch für die Kritik ihres Inhalts von nicht zu unterschätzender Bedeutung ist. Über diese Vorzüge darf aber nicht vergessen werden, dass die Konstruktion eines Experimentellen Fragebogens ebenso wie die Interpretation seiner Meßergebnisse eingehende Kenntnisse über den Untersuchungsgegenstand voraussetzt, beides also den gedanken- und theorieleeren Einsatz in der Forschung ausschließt.

Als ein Beispiel für einen Experimentellen Fragebogen ist in Tabelle 4 das dreifaktorielle Design des [“Moralische Kompetenz-Tests” (MKT, Lind 2019)] dargestellt. Es werden Handlungsdilemmata vorgestellt und anschließend Argumente für und gegen das Verhalten des Protagonisten zur Beurteilung vorgelegt. Aufgrund des dreifaktoriellen Designs repräsentiert jedes Argument je eine bestimmte Ausprägung dreier Bewertungsdimensionen: einen bestimmten Dilemma-Kontext, eine bestimmte Urteilsqualität sowie eine bestimmte Meinung zur Lösung des jeweiligen Dilemmas. Dieses Fragebogenexperiment basiert auf Kohlbergs (1969) Theorie der kognitiv-moralischen Entwicklung und auf Annahmen über das Antwortverhalten in der Befragungssituation (vgl. Lind 1980c). Die drei wesentlichen Annahmen sind, (a) dass die Bewertung von Argumenten für und gegen eine bestimmte Lösung eines moralischen Dilemmas von der “moralischen” Qualität des Arguments mitbestimmt ist, genauer gesagt durch den Entwicklungsstand des moralischen Bewusstseins einer bestimmten Person und ihrer Wahrnehmung des Arguments als moralisches; (b) dass die Bewertung von Argumenten auch (und in früheren Entwicklungsstadien vorwiegend) davon abhängt, ob diese Argumente mit der eigenen Meinung zur Lösung des Dilemmas übereinstimmen oder nicht (Meinungsübereinstimmung, ‘opinion-agreement’); und (c) dass die Bewertung von Argumenten differenziert erfolgt, d.h. dass sich darin eine Koordination des moralischen Bewusstseins mit dem Engagement für eine bestimmte Meinung und der Berücksichtigung des jeweiligen Problem-Kontextes spiegelt. Wir vermuten also, dass bestimmte Konstellationen dreier Faktoren das Urteilsverhalten in (so definierten) moralischen Situationen beeinflussen. Wir erwarten nicht, dass wir damit das konkrete, historisch einmalige Verhalten jeder Person vorhersagen können – das wäre auch in Bezug auf andere Verfahren keine realistische Erwartung – sondern vielmehr, dass sich mit Hilfe dieser Erfassungsmethode eine größere kategoriale Überlappung zwischen der Realität und dem theoretischen System ergibt als dies in der Psychometrie bislang der Fall war.

Tabelle 4
Experimenteller Fragebogen:
Das Design des [Moralische Kompetenz-Tests (MKT)]

Unabhängige Variablen (Faktoren)			Abhängige Variablen (Verhalten)								
1. Geschichte (Kontext, Dilemma) ^a	2. Moralische Qualität des Arguments (Stufe) ^b	3. Meinungsübereinstimmung (Pro-Con)	Beurteilung der Akzeptabilität von vorgegebenen Argumenten								
Diebstahl (Arbeiter)	1	Pro	Ich lehne völlig ab				Ich akzeptiere völlig				
		Con	-4	-3	-2	-1	0	+1	+2	+3	+4
	2	“	-4	-3	-2	-1	0	+1	+2	+3	+4
	3	“	-4	-3	-2	-1	0	+1	+2	+3	+4
	4	“	-4	-3	-2	-1	0	+1	+2	+3	+4
	5	“	-4	-3	-2	-1	0	+1	+2	+3	+4
Sterbehilfe (Arzt)	1	“	-4	-3	-2	-1	0	+1	+2	+3	+4
		“	-4	-3	-2	-1	0	+1	+2	+3	+4
	2	“	-4	-3	-2	-1	0	+1	+2	+3	+4
	3	“	-4	-3	-2	-1	0	+1	+2	+3	+4
	4	“	-4	-3	-2	-1	0	+1	+2	+3	+4
	5	“	-4	-3	-2	-1	0	+1	+2	+3	+4
6	“	-4	-3	-2	-1	0	+1	+2	+3	+4	

^a Die in unseren Untersuchungen benutzte Version enthält zwei Dilemmata mit unterschiedlichem ‘moralischem’ Anforderungsgehalt. Erweiterungen sind möglich und je nach Forschungsinteresse sogar wünschenswert (vgl. Lind & Wakenhut 1980; [Lind 2019]).

^b Die Qualitätsunterschiede der Argumente orientieren sich an den sechs Kohlbergschen Stufen in der ursprünglichen Fassung (Kohlberg 1969). Zwar erschwert dies den Vergleich mit Untersuchungen, die die neuere Version des Stufenschemas benutzen. Wir sind jedoch der Ansicht, dass die Revision des Schemas weder notwendig war noch wesentlich zur Verbesserung der Theorie beigetragen hat. Die beobachteten Fälle von Entwicklungsregression widerlegen nur die starke Form der Entwicklungslogik. Die schwächere Version, die nur Invarianz der Sequenz annimmt, konnte bereits mit dem alten Schema empirisch überzeugend belegt werden.

Belege für die Nützlichkeit des Ansatzes

Das Konzept des Experimentellen Fragebogens ist noch neu und kann deshalb erst sektoral evaluiert werden. Immerhin liegen aus dem Projekt *Hochschulsozialisation* (Lind 1978a, 1980d, f; [Lind 2002]) sowie aus anderen Forschungsprojekten schon Erfahrungen mit dem [Moralische Kompetenz-Test] und ähnlichen Verfahren vor (Brügelmann 1979; Cattepoel 1977; Krämer-Badoni & Wakenhut 1978; Portele 1979, [Lind 2002, 2019]), so dass eine gewisse Bewertung der Nützlichkeit dieses Konzepts schon jetzt möglich ist.

Zwei Befunde scheinen mir dafür besonders bedeutungsvoll:

(1.) Obgleich der Test im Design und in der Auswertung völlig anders konzipiert ist als Kohlbergs originäre und hochgradig elaborierte Interviewmethode haben sich einige seiner grundlegenden Befunde auch mit unserem Verfahren bestätigen lassen (Lind 1980d). Das belegt sowohl die Methodenunabhängigkeit der Kohlbergschen Theorie als auch die Nützlichkeit des [“Moralische Kompetenz-Tests”] für die Erfassung [der Moralkompetenz].²

(2.) Im Vergleich zu vielen herkömmlichen Verfahren der Persönlichkeitsdiagnostik zeigt dieser Experimentelle Fragebogen eine beachtliche kategoriale Überlappung. Während jene selten mehr als 10 Prozent der Antwortvarianz “erklären” können, ist es mit dieser Methode ein wesentlich größerer Prozentteil. Unterstellt man, dass nur das moralische Bewusstsein für das Urteilsverhalten maßgebend ist, dann können wir damit bereits durchschnittlich 27 Prozent der individuellen Antwortvarianz erklären. Dabei spricht die große Streuung der Verhaltensrelevanz des moralischen Bewusstseins eindeutig gegen das Gemeinsame-Eigenschaften-Modell der klassischen Psychometrie!

Nimmt man an, dass das Urteilsverhalten von erwachsenen Personen differenzierter und meist das Resultat einer dynamischen Koordination von mehreren Faktoren ist (in diesem Fall von dem moralischen Bewusstsein, von dem Engagement für eine bestimmte Meinung zum konkreten Fall und von der Art des Dilemmas), dann lassen sich durchschnittlich über 60 Prozent des individuellen Urteilsverhaltens theoretisch rekonstruieren.

² Früher: “Moralisches Urteil-Test” und “moralisches Bewusstsein”. Wie ich später (Lind 2019) darlege, misst der MKT keine bewussten Prozesse wie verbalisiert Urteile, sondern einen Aspekt (moralische Kompetenz), der sich im Muster des Antwortverhaltens manifestiert. [Nachtrag 2019; GL]

Unsere bisherigen Erfahrungen mit dem integrationistischen Ansatz, speziell mit dem hieraus entwickelten Konzept des Experimentellen Fragebogens, sind alles in allem ermutigend. Zwar hat uns nicht zuletzt dieser Versuch gewahr gemacht, wie schwierig es ist, dem Ideal der Umsetzung einer Theorie in eine adäquate Forschungsmethode näher zu kommen: Aber es hat sich auch gezeigt, dass weder die a-theoretische Verfeinerung der Psychometrie noch der methodendogmatische Standpunkt zu Fortschritten in der Hochschulsozialisationsforschung führen. Vielmehr scheint ein *methodenrationaler* Standpunkt erfolgversprechend, der bestimmte Methoden nicht a priori bestimmten Gegenstandsbereichen zuordnet sondern sich dem Problem der Adäquatheit, d.h. der *theoretischen Validität* stellt. Die Abstimmung zwischen Theorie, Methode und empirischen Befunden ist als ein regelkreisartig verbundener, hermeneutisch-zirkulärer Prozeß zu verstehen, der in Forschung und Praxis iterierend durchlaufen wird. Es ist der Standpunkt, der bereits von Wilhelm Dilthey und Charles S. Pierce vertreten wurde und der heute u.a. von Peter Krausser (1968), dessen ausgezeichnete Rekonstruktion der Diltheyschen Handlungs- und Wissenschaftstheorie in vielfacher Hinsicht als Ergänzung und Vertiefung zu meinem Referat gelesen werden kann, und von Lawrence Kohlberg (1979) fortgeführt wird. Gerade Kohlberg gewinnt dadurch über seinen inhaltlichen Beitrag zur Sozialisationsforschung hinaus auch eine methodologische Bedeutung, die bislang noch zu wenig gewürdigt wurde.

Dieser methodenrationale Standpunkt ist eingebunden in die Vorstellung, dass, wie Krausser (1968) sagt, der "Erkenntnisprozeß als Prozeß des Lernens, der Verbesserung nicht nur unserer Aussagen, sondern auch unserer Fragestellungen (Ziele), Kriterien (Werte) und Methoden (Regeln, Handlungsweisen)" gilt (S. 190). Wir können uns nicht auf den Standpunkt stellen, dass einer dieser Bestandteile absolute Gültigkeit besitzt und nicht durch weitere Forschung in Frage gestellt werden könnte. Deshalb ist auch wünschenswert, dass Theorien der Sozialisation in der Hochschule soweit expliziert werden, dass ihre empirischen Implikationen hervortreten und damit der Überprüfung zugänglich werden. Dies ist der Punkt an dem meine Ausführungen sich mit den Schlußüberlegungen von Reichwein (1980) treffen. In der Tat, man "müßte entweder einen direkteren Weg (als den über hypothetische Annahmen) finden, (um die Identitätsformen zu erfassen), oder die empirische Brauchbarkeit der interaktionistischen Identitätstheorie doch wieder infragestellen" (S. 25; Zusätze in Klammern von mir). Vielleicht öffnet der hier vorgestellte Ansatz eine adäquate Möglichkeit zur Entscheidung dieser Frage.

Kapitel 2

KOGNITIV-STRUKTURELLE ANALYSEN VON POLITISCHEN
EINSTELLUNGSSYSTEMEN: ÜBERLEGUNGEN UND BEFUNDE AUS
DER FORSCHUNGSGRUPPE *HOCHSCHULSOZIALISATION*¹*Einleitung*

Kognition und Struktur sind in der Sozialisationsforschung gewiß keine unbekanntenen Wörter – mit einigem Recht werden sie hie und da bereits als Modewörter abgetan. Wir meinen jedoch, dass sie recht begriffen, der Analyse von Einstellungssystemen eine noch zu wenig genutzte Perspektive eröffnen. Wie kaum ein anderer hat Jean Piaget dazu beigetragen, die Begriffe Kognition und Struktur theoretisch fruchtbar zu machen und ihnen empirische Bedeutung zu verleihen. Sein Buch “Das moralische Urteil beim Kinde” (1973/1932) wird wohl zu Recht als ein epochemachendes Werk der politischen Sozialisationsforschung angesehen (Harten 1977). Was seine Arbeit besonders wichtig für die politische Psychologie macht, ist, dass in ihr die kognitiv- strukturelle Komponente des im weitesten Sinne moralischen Bewusstseins unmittelbar mit der Möglichkeit der demokratischen Vergesellschaftung des Menschen in Verbindung gebracht wird. Für Piaget ist die “intellektuelle Dezentrierung, das Festhalten an ethischen Vorstellungen auch wenn sie den momentanen Interessen der eigenen Person widersprechen, eine notwendige Voraussetzung für demokratische Zusammenarbeit.

Der große Verdienst Piagets wird gewiß nicht geschmälert, wenn man feststellt, dass eine Reihe von unklaren und widersprüchlichen Aussagen zur Kritik herausfordern (man denke nur an die Korrektur seiner Theorie der Sprachentwicklung durch Wygotzki oder die Korrektur seiner Theorie der Moralentwicklung durch Kohlberg), und dass andere Ansätze, wie die Psychoanalyse, die Einstellungstheorie oder der Symbolische Interaktionismus je spezifische Verdienste haben, die durch die kognitive Entwicklungstheorie Piagets nicht überflüssig werden.

Ich will im Folgenden von dem Versuch einer Integration dieser verschiedenen Ansätze berichten, den wir im Forschungsprojekt *Hochschulsozialisation* in Konstanz unternommen haben, um Fragestellungen, Untersuchungsanlage und Auswertungsstrategien im Bereich der universitären Sozialisation durch ein das Gemeinsame betonende Grundmodell zu koordinieren.

Wir haben weder eine fertige Theorie zu bieten, noch sind wir in der Lage, die empirische Gültigkeit unseres Grundmodells der moralpolitischen Persönlichkeit umfassend zu belegen. Aber schon jetzt scheint es möglich, einige Probleme der Sozialisationsforschung damit auf den Begriff zu bringen und Konsequenzen für eine adäquate Forschungsstrategie zu ziehen. Auch können wir erste empirische Überprüfungen vorlegen, die das Vorgehen und die Nützlichkeit kognitiv-struktureller Analysen illustrieren.

Die Unabgeschlossenheit der projektinternen Diskussion wie auch der Umstand, dass unsere Längsschnittuntersuchungen zur Hochschulsozialisation² noch im Gange sind, lassen es sinnvoll erscheinen, unsere Überlegungen und Befunde in Thesenform zu bilanzieren. Ich werde dabei von der Möglichkeit Gebrauch machen, auf einen ausführlicheren Aufsatz hinzuweisen, den ich zusammen mit meinen Kollegen Johann-Ulrich Sandberger und Tino Bargel hierzu verfaßt habe.³ Ich will die Gelegenheit dieses Vortrages auch dazu nutzen, einiges, was dort nur angedeutet ist, hier pointierter zu fassen.

1. Gegen einen soziologistischen Reduktionismus

Meine erste These greift Piagets Kritik an Durkheim auf. Sie wendet sich gegen eine soziologistische Reduktion der Rolle der moralisch-politischen Erkenntnis in der Demokratie, die das Individuum auf die Übernahme des 'allgemeinen' Willens verweist und deshalb der kognitiv-strukturellen Verankerung demokratischer Werthaltungen keine weitere Funktion zubilligt als die der Normenkenntnis. Piaget verkennt nicht, dass die das individuelle Bewusstsein füllenden Werthaltungen und Vorstellungen sozial vermittelt und sozialen Ursprungs sind (er will also auch keine psychologistische Reduktion), aber er wendet sich gegen die Durkheimsche Theorie, wonach allein die Gesellschaft mittels ihrer wissenschaftlichen Einrichtungen zur rationalen Rekonstruktion der "objektiven moralischen Wirklichkeit" (Durkheim 1976/1924, S. 90; vgl. auch S. 98, 116, 122 und 135) in der Lage ist. Max Webers Hinweis auf die Probleme der kollektiven Rationalität, der Bürokratisierung, und ebenso die hierzulande gemachte Erfahrung, wie die (Sozial-)Wissenschaft durch die Naziherrschaft ausgeschaltet oder korrumpiert wurde,

machen nachdrücklich deutlich, dass das Programm der Aufklärung, der mündige Staatsbürger, durch keine noch so demokratisch organisierte Moralwissenschaft ersetzt werden kann. Schluchter (1979) hat kürzlich in einem Essay über den westlichen Rationalismus von der Notwendigkeit einer Balance zwischen "formaler" (kollektiver bzw. bürokratischer) Rationalität und "materieller" (individueller) Rationalität gesprochen. Vielleicht wird es möglich sein, genauer anzugeben, wie eine solche "Balance" konkret aussieht und wie sie in einer "totalen Rationalität" (Lukacs) aufgehoben werden könnte, wenn die Interaktion zwischen Individuum und Gesellschaft dort, wo sich kollektive und individuelle Rationalität überschneiden, genauer bekannt ist. Dies ist bei Akademikern nicht nur faktisch wie bei jedem gesellschaftlichen Wesen der Fall, sondern oft auch bewusstseinsmäßig reflektierend. Gerade bei Akademikern wird die formale Rationalität des Staates infolge ihrer besonderen sozialen Position zu einem Teil ihrer eigenen materiellen Rationalität (die zuge dachte größere Verantwortlichkeit wird, wie Bargel et al. 1973 zeigten, von Akademikern mitgenutzt zur Reklamation größerer Gratifikationen). Umgekehrt wird die materielle Rationalität, der Entwicklungsstand des individuellen Moralbewusstseins, zum wichtigen Moment der Rationalität, die dem formalen Konstrukt 'Staat' möglich ist. Position und Charakter befinden sich also in einem dialektischen Verhältnis. Was das eine zulässt, wird dem anderen möglich; was das eine fordert, wird dem anderen zur Aufgabe. Die Sozialisation von Akademikern zeigt deshalb am deutlichsten an, ob die 'Balance' zwischen individueller und kollektiver Rationalität in der Demokratie gelingt. Nicht zufällig wird gerade von Akademikern in unserer Gesellschaft kritische Rationalität und autonome Urteilsfähigkeit erwartet, wie unsere Umfragen belegen (Bargel & Framhein 1976; Sandberger & Lind 1979).

Weitergehende Modelle der Entwicklung einer das Individuum und die Gesellschaft umfassenden Rationalität versuchen, die Balance inhaltlich als neue Qualität zu bestimmen. So Kants Idee einer vollkommenen, nach Regeln der Gerechtigkeit regierten Republik", die in Rawls (1975) Theorie der Gerechtigkeit, Kohlbergs (et al. 1974) Just-Community-Approach und Höffes (1975) Strategien der Humanität ihre moderne Interpretation gefunden hat. So auch die Idee der kommunistischen Gesellschaft, wie sie von Bloch oder Bahro weitergedacht wird. Solche Ideen werden approximiert in den Verfassungen von Staaten und in dem postkonventionellen moralischen Bewusstsein ihrer Bürger. In welchem Maße sie als Entwicklungsnorm eingelöst werden, ist die Frage der 'praktischen Anthropologie' (Kant). Die Thesen zwei bis vier beschränken sich auf die psychologische Problemstellung, nämlich auf den

Versuch, die kognitive Verankerung ethischer Werthaltungen durch strukturelle Analysen nachzuweisen.

2. Von der Nichtseparierbarkeit von Struktur und Inhalt

Die These zwei handelt von dem grundlegenden Problem, Struktur und Inhalt in Beziehung zu setzen. Sie besagt, dass die kognitiv-strukturelle Komponente des Bewusstseins nur analytisch unterscheidbar, aber nicht dinglich trennbar ist von den Bewusstseinsinhalten. Nicht wenige Schwierigkeiten der Sozialisationsforschung resultieren aus der Mißachtung dieses Nichtseparierbarkeits- Postulats.

Soll, wer 'dogmatisch' an der Idee der Demokratie festhält, als 'strukturell gleich' angesehen werden wie der, der 'dogmatisch' an seiner Meinung zu einer konkreten Sachlage festhält? Kann, wer differenziert urteilt, deshalb als inkonsistent bezeichnet werden? Ich glaube, man kann die erste Frage unter Hinweis auf den Inhalt der dogmatischen Struktur verneinen, und die zweite Frage unter Hinweis auf ihre inhaltliche Unbestimmtheit als unbeantwortbar zurückweisen. Man kann danach fragen, nach welchen Kriterien differenziert geurteilt wurde und in Bezug auf welche Motivation eine Handlung inkonsistent ist. Wenn wir von kognitiv-strukturellen Analysen sprechen, dann meinen wir also nicht etwas dinglich von Inhalten Getrenntes, wie in vielen Sozialisationsstudien offensichtlich für möglich gehalten (vgl. u.a. Ijzendoorn 1979), sondern die relationale Struktur zwischen Inhalten, z.B. zwischen verschiedenen Werthaltungen und zwischen Werthaltungen und funktionalen Verhaltensweisen. Es würde zu weit führen, hier unser Grundmodell einer dynamisch-strukturellen Einstellungs-Einheit ('trait') aus den verschiedenen Ansätzen, wie Psychoanalyse, Einstellungstheorie, Persönlichkeitstheorie und kognitiver Entwicklungstheorie abzuleiten – das wurde skizzenhaft an anderer Stelle unternommen (Lind, Sandberger & Bargel 1980; Lind 1980c).

Es mag für das Verständnis unseres Ansatzes jedoch hilfreich sein, sich dieses Grundmodell schematisch vor Augen zu halten (vgl. Abbildung 1). Vor allem ist wichtig, die Beziehungsstruktur als zentrales theoretisches Bestimmungsstück der individuellen Persönlichkeit zu begreifen. In ihr spiegelt sich die sprachlich-kognitiv festgelegte Definition der Wirklichkeit, d.h. das System der sogenannten *a priori* Kategorien, die die Interaktion mit der Umwelt bestimmen.

Diese a priori Kategorien brauchen allerdings nicht als angeboren, d.h. jeder Erfahrung vorgängig angesehen werden, sondern unterliegen vermutlich, wie bereits Dilthey argumentiert und Piaget gezeigt hat, ihrerseits der Kritik durch Erfahrung und konkurrierende Kategorien. Kognitiv-strukturelle Analysen in der Sozialisationsforschung sind also gerichtet auf den Prozeß der Strukturierung des Bewusstseins nach bestimmten dominanten Kategorien (Werthaltungen, moralischen Prinzipien, politischen Ideologien) und der Organisation von Handlungen unter diesen Kategorien. Es geht, wie Max Weber es formulierte, um die "Konstanz der inneren Verhältnisse der Persönlichkeit zu bestimmten letzten Werten und Lebensbedeutungen. Erst durch die von der Ethik aufgegebenen Konstanz wird wechselseitiges Verstehen und innere Orientierung möglich. Unter inhaltlicher Perspektive tritt dazu noch die Frage, welche Richtung oder Ausprägung dieser Kategorie jeweils konkret vorhanden ist, ob z.B. der Wert "soziale Gleichheit" befürwortet oder abgelehnt wird.

Methodische Konsequenzen des kognitiv-strukturellen Ansatzes

Das soeben behandelte Modell hat trotz, oder vielleicht wegen seiner Einfachheit und der damit verbundenen Transparenz, eine Reihe von theoretischen und methodischen Implikationen. Ich kann hier nicht alle diskutieren. Die mir 25 am wichtigsten erscheinende Konsequenz will ich in eine dritte These fassen. Sie besagt, dass das von der klassischen Test- und Fragebogentheorie aufgestellte Kriterium der Reliabilität substantielle Annahmen über die Natur der Persönlichkeit enthält, die mit dynamisch-strukturellen Modellen nicht vereinbar sind.

Ein allzu lascher Umgang mit dem Konzept der Reliabilität, das in der mathematischen Fehlertheorie präzise als Ergebnis einer reinen Zufallsauswahl bzw. eines unverzerrten Zufallsprozesses definiert ist, macht es oft schwer, diesen Widerspruch zu erkennen und ihn als empirische Frage einem Entscheidungsexperiment zuzuführen. Hier bahnt sich jetzt eine Wende an. Lumsden (1976) sowie deutsche Testtheoretiker (Kempf 1978) haben die semantischen Probleme der probabilistischen Testtheorie im Bereich der Fähigkeitsdiagnostik aufgewiesen. Wakenhut hat ähnliche Überlegungen in Zusammenhang mit der Skalierung des moralischen Urteilsverhaltens angestellt (Krämer-Badoni & Wakenhut 1978). In unserem Bestreben, es dem physikalischen Messen gleichzutun, haben wir in der Sozialisationsforschung meist großen Wert auf "reliable" Skalen gelegt und uns

damit, wie sich jetzt herausstellen könnte, in Widerspruch zu den eigenen Modellen der Persönlichkeit begeben. Auf die damit verbundenen Probleme in der politischen Psychologie haben bereits Converse (1970) mit dem Begriff der non-attitudes sowie Hartmann & Wakenhut (1973) anlässlich ihrer Beobachtung eines Bedeutungswandels von politischen Kategorien aufmerksam gemacht. Die Frage wurde erneut im Zusammenhang mit dem Aspekt der Zentralität von politischen Einstellungen aufgeworfen (Lippert & Wakenhut 1979). Der klassische Ansatz der Psychometrie wurde häufig kritisiert unter Heranziehung seines Entstehungs- und Verwendungszusammenhangs, ohne dass diese Kritik letztlich überzeugen konnte. Die oben vorgetragene Kritik, dass er etwas anderes realisiere als er zu realisieren beabsichtigt, trifft diesen Ansatz jedoch härter. Die versteckte, substantielle Annahme, die sich mit ihm verbindet, kann expliziert und empirisch falsifiziert werden.

Die Fehlertheorie geht davon aus, dass alle inkonsistenten Antworten in einem Gruppenexperiment dem Erhebungsinstrument zuzuschreiben sind, also die Größe der Inkonsistenz nicht durch Eigenschaften der gemessenen Personen bestimmt ist. (Dies ist eine etwas vergrößernde Schlußfolgerung, da bei begrenzten Skalen und ungleicher Meßwertverteilung bereits aus logischen Gründen keine Unabhängigkeit des Meßfehlers mehr gegeben ist, ein Hinweis, den ich Wilhelm Kempf verdanke). Wie wir an anderer Stelle (Lind 1980b) gezeigt haben, läßt sich diese Annahme, um es vorsichtig auszudrücken, nicht als generell gültige aufrechterhalten. Der Grad der Konsistenz ist abhängig von der Person des Befragten und von der entwicklungsmäßigen "Schwierigkeit" der zu beurteilenden Stellungnahme.

Entwicklung als Integration und Differenzierung

Die Gegenposition will ich in meiner vierten These formulieren. Sie besagt, dass sich eine entwickelnde Persönlichkeit immer gleichzeitig in Phasen des Äquilibrium und des Desäquilibrium befindet, sich also immer in Bezug auf manche regulative Ideen relativ konsistent, in Bezug auf andere aber relativ inkonsistent verhält. Sie besagt weiter, dass es Kernkonzepte gibt, die typisch für bestimmte Entwicklungsphasen sind und ihrerseits in einem Subsumptionsverhältnis stehen und damit der Ontogenese einen gewissen invarianten Charakter verleihen. Das bekannteste Beispiel hierfür ist wohl Kohlbergs Modell der kognitiv-moralischen Entwicklung, dessen Stufen durch sechs soziale Perspektiven markiert werden, unter denen das Individuum jeweils den gerechten Ausgleich von Interessen sucht.

Es gibt Hinweise dafür, dass sich diese Grundsequenz der Integration und Differenzierung ebenso bei politischen, sozialen, beruflichen oder anderen Einstellungen finden läßt (die Arbeiten von Adelson & O'Neill 1966, von Hess & Torney 1967 sowie von Newcomb 1943 sind hierfür gute Beispiele), und dass deshalb Inkonsistenzen des Verhaltens in Bezug auf diese Einstellungen nicht allein als Meßfehler interpretiert werden können.

Was diese These besagt, läßt sich nicht in einem Experiment ein für allemal beweisen, sondern muß für jeden Untersuchungsgegenstand aufs neue präzisiert und empirisch überprüft werden. Wenngleich schon jetzt feststeht, dass die Annahmen der klassischen Psychometrie nicht mehr als allgemein gültig gelten können, bleibt die Interpretation von Inkonsistenzen als Meßfehler jedoch weiter als Möglichkeit erhalten.

Wir stehen im Projekt *Hochschulsozialisation* erst am Anfang der Realisation des in diesen vier Thesen angelegten Forschungsprogramms. Wir haben als Reaktion auf den Mangel adäquater Erhebungskonzepte verschiedene Strategien ausgearbeitet, um den kognitiv-strukturellen Aspekt von Einstellungssystemen zu analysieren, d.h. um neben der Frage nach der Ausprägung auch die Frage nach dem individuell verschiedenen Grad des Vorhandenseins und der Organisation von dynamisch-strukturellen Persönlichkeitsmerkmalen zu beantworten.

Am weitesten gediehen sind die Analysen der Struktur von Gesellschaftsbildern, bei der wir erstaunlich hohe interkulturelle Invarianzen haben feststellen können (Sandberger 1980), der kognitiven Organisation von soziopolitischen Werthaltungen (Sandberger 1979), sowie der Struktur des moralischen Bewusstseins, die wir mit dem neuen, oben bereits erwähnten Konzept des Experimentellen Fragebogens einer detaillierten Analyse unterzogen haben. Das Rationale, das Design und verschiedene neue Möglichkeiten der Auswertung von Experimentellen Fragebogen sowie ihrer Gütekontrolle habe ich an anderer Stelle ausführlicher behandelt (Lind 1980c).

Zusammenfassung

Zusammenfassend ist zu sagen, dass sich der mit den vier Thesen markierte Ansatz bereits gut bewährt hat. Es ist dadurch nicht nur möglich, bestimmte Beschränkungen bisheriger Sozialisationsforschung konstruktiv zu überwinden, sondern auch, wichtige Annahmen und Befunde von sogenannten 'qualitativen' Forschungsansätzen auf eine breitere empirische Basis zu stellen. Ich möchte nur darauf hinweisen, dass sich einige zentrale Annahmen Kohlbergs über die Natur des moralischen Bewusstseins auf überraschend deutliche Weise haben bestätigen lassen (Lind 1980d): Die Stufen der Moralentwicklung zeigen die vorhergesagte Präferenzordnung und Quasi-Simplex-Struktur der Interkorrelationen; auch erweist sich die bislang empirisch noch nicht direkt überprüfte Annahme einer gewissen Parallelität von inhaltlicher und struktureller Komponente des moralischen Bewusstseins als empirisch stichhaltig (vgl. Abbildung 2, erste Zeile)⁴; schließlich gibt es erste Hinweise darauf, dass Studienanfänger in der Tat erst Stufe drei kognitiv voll integriert haben. Dieses letzte Ergebnis bedarf allerdings noch der Bestätigung durch intra-individuelle Auswertungsverfahren, an denen wir gerade arbeiten. Von allgemeiner Bedeutung ist weiterhin, dass sich die bislang als Meßfehler gedeuteten Inkonsistenzen als substantielle Komponenten der individuellen Persönlichkeit erweisen, und dass diese kognitiv-strukturelle Komponente systematisch mit demokratischer Orientierung und Ich-Stärke kovariiert (vgl. Abbildung 2).

Liegt eine demokratische Werthaltung vor, dann zeigt sich in der Struktur des moralischen Urteilsverhaltens ein stärkeres Maß an “intellektueller Dezentrierung (Piaget) und Konstanz der Verhältnisse” (Weber) in Bezug auf ethische Orientierungen. Ich meine, dies zeigt deutlich, dass den kognitiv-strukturellen Analysen nicht nur mit Recht eine größere Bedeutung für die Sozialisationstheorie zukommt, sondern dass sie auch sinnvoll und fruchtbar in der empirischen Forschung eingesetzt werden können.

Anmerkungen

- 1) Beitrag zu dem round-table “Politische Psychologie” vom 25.9. bis 27.9.80, veranstaltet vom Zentrum für Umfragen, Methoden und Analysen (ZUMA), Mannheim.
- 2) Vgl. Forschungsgruppe *Hochschulsozialisation* (1979).
- 3) Lind, Sandberger & Bargel 1980.
- 4) Vgl. Lind, Sandberger & Bargel 1980; die strukturelle Komponente des moralischen Bewusstseins läßt sich bei diesen Akzeptabilitätsprofilen aus der Prägnanz ihrer stufentypischen Differenzierung ablesen.

Nachtrag [2019]

Auch heute, fast dreißig Jahre nach Erstellung dieses Aufsatzes ist sein Inhalt hochaktuell - aus eher bedauerlichem Anlass: die sozialwissenschaftliche und psychologische Forschung basiert heute immer noch auf einer rein statistischen Testtheorie, die psychologische Tatsachen weitgehend ignoriert. Dabei haben wir gezeigt, dass eine neue, psychologische Testtheorie die Tür für eine progressive Wissenschaft öffnet. Mit dem Moralische Kompetenz-Test (MKT), der früher anders hieß (Moralisches Urteil-Test, MUT), konnten wir wichtige Annahmen der kognitiven Entwicklungstheorie der Moral von Piaget und Kohlberg erstmals objektiv testen. Die MKT-Forschung hat sie teilweise eindeutig bestätigt, wie Annahme von zwei verschiedenen, aber untrennbaren Aspekten der Moral, besonders die Existenz des kognitiven Aspekts der *Moralkompetenz* (Lind 2019). Teilweise konnten wir damit auch Annahmen eindeutig widerlegen, wie Kohlbergs Postulat der invarianten Sequenz, also der Nicht-Regression.

.

Kapitel 3

RELIABILITY OF TEST OR STRUCTURE OF PERSONALITY?
EMPIRICAL FINDINGS ON THE PROBLEM OF INTERPRETING
JUDGMENT CONSISTENCY.*

When evaluating the data of the FORM-project we are frequently confronted with the dilemma of how to interpret response consistency. On one hand we conceptualize students' personality development as a complex reorganization of attitudes and behaviors. When we say that through the process of socialization the student acquires new concepts which guide his actions then we think of him, or her, to grow in the degree of consistency with which this concept becomes gradually applied to behavioral decisions or judgments. That is, we conceive of consistency as an aspect of the students' attitude structure (A. Lind 1978b; Bromberek, 1978; also the papers in Framhein, 1980).

On the other hand, in constructing attitude scales we utilize the classical concept of attitude measurement which defines attitude as an "affective tendency toward a social object" (Thurstone, 1931) and also that with regard to this tendency the individual respondent is consistent. According to the classical attitude concept on which most of our standard methods of scale construction are based any inconsistency of responses is attributed to the scale used. One speaks solely of unreliability of the measure and not of inconsistency of individual behavior. Hence the assumption that consistency is a sign of the individual's personality structure contradicts this presupposition.

If we want to assess not so much the surface change of affective tendencies but the development of the underlying cognitive organization of attitudes (their "deep structure") then we must attempt to resolve the dilemma between methodology and theory.

So far, most of the corroborating evidence for the utility and adequacy of the developmental approach comes from so-called qualitative (interview) research (cf. Kohlberg, 1969; Perry, 1970; and also the research of the Swedish FORM-participants).

But there is only few quantitative research into the development of attitude structure which could provide evidence for either thesis. Since this is a crucial question for the FORM-project I did some analyses which should help to clarify this issue using data from the Abiturienten-untersuchung of the Konstanz-team. I confined my investigation to the *Moral Compezenze Test* (MCT; Lind, 1980e) though, in principle, any other scale of the FORM-questionnaire might have been taken. The moral-judgment-test was preferred, however, as it opens up the possibility of testing the hypothesis that response consistency is not only a personal characteristic but also indicates the level of personality development attained. Thus the three hypotheses to be tested are:

1. Inconsistency is due to the test; it is independent of the person measured (= classical test theory).
2. Inconsistency is due to personal differences with regard to their degree of personality integration (= structure of personality).
3. Inconsistency indicates the level of personality integration, i.e., a developmental trend which can be expected on the basis of cognitive-developmental theory.

In order to understand the results of that analysis it may be helpful to recollect the design of the [MCT] and some intermediate steps of data evaluation first. In Table 1, the “anatomy” of an individual student’s pattern of moral judgment behavior is depicted. It clearly indicates that this person rejects the lower stage arguments and rather accepts the higher stage arguments presented to him (note that in the German version two stories are used instead of only one as in the international core of the FORM-questionnaire). In computing the sums, or the averages, for each of the six stages we still follow classical methods of attitude measurements. The average scores tell us something about this individual’s attitudes toward each of the six stages of moral judgment. Graph 1 gives the combined attitude profiles of students in the 1st semester and 9th semester for further comparison across the participating countries.

In addition to this we have, however, gone beyond classical methodology and computed scores which reflect the individual’s consistency of judgment with regard to several dimensions of judgment (cf. Table 1, lower part). The ways of analysis and the interpretation of these scores have been discussed in two other papers (Lind 1978a, 1980g). Here, we need only to attend to

that number which gives us the amount of consistency with regard to the dimension which the test is to measure primarily, namely with regard to factor Stage. The respective figure says that this student's responses are 63 per cent consistent with regard to the Stage-dimension built into the test. That is, within this individual the factor Stage correlates as high as $r = 0.79$ with the individual's rating behavior.

In classical view the remaining inconsistency is due to measurement error introduced by the test. If this hypothesis were right then we should expect that our instrument measures equally well, or equally poorly, in

Table 1

Anatomy of an Individual Pattern of Moral Judgment Behavior (N = 1)

a) Measures of moral attitudes (affective aspect)

		Mercy Killing (Doctor)			Theft (Workers)			
		PRO	CON	\bar{x}	PRO	CON	\bar{x}	total \bar{x}
Stage	1	-4	-4	-4	-4	-4	-4	-4
	2	-4	-4	-4	-4	-4	-4	-4
	3	-4	-4	-4	-4	-4	-4	-4
	4	-4	2	-2	-4	3	-0,5	-0,7
	5	-4	3	-0,5	-4	4	0	-0,2
	6	3	4	3,5	3	3	3	3

b) Some Measures of Moral Judgment Structure

Modal Preference Stage: Stage 6

Proportion of Variance accounted for by factor ...

- STORY 0 %

- PRO-CON * 12 %

- STAGE 63 % [Today named "C-score"]

* Corrected for general opinion about the moral problem

all instances, that is, regardless of the person being measured. This is without doubt the implication of all approaches to attitude scale construction that rest on the classical paradigm advanced by scholars like Thurstone and Likert. The developmental paradigm forwarded by, e.g., Kohlberg (1969) and Perry (1970) as well as the attitude structure paradigm (G. Converse, 1970), would, at least implicitly, forecast that the individuals will greatly differ with regard to response consistency, i.e., that the individuals and not the assessment instrument accounts for “measurement error”. Looking at Graph 2 we indeed see that, in our data, the latter hypothesis is more likely to be correct than the former. The solid line indicates that the degree of consistency with regard to factor Stage varies greatly across our 708 high school graduates. There is a group of considerable size which responds totally inconsistently; and there is also a large number of students the responses of whom are almost perfectly consistent. The dotted line reveals the distribution of students with regard to factor *Pro-Contra*, i.e. with regard to the degree that the moral argument presented agrees or disagrees with the respondent’s own opinion about the solution of the moral problem. It shows that many students have also (or solely) attended to this factor when evaluating the scale items of the [MCT]. They have chosen a focus of concern different from that which the constructor of the test primarily had in mind. What are the consequences of these findings? First, I believe, we should not conceive of measures of consistency as being of mere technical importance anymore. We have good reasons now to regard these measures as meaningful indicators of substantial aspects of students’ personality structure. We have, in addition, found substantial correlations between such consistency measures and other features of students’ personality (e.g. ego-strength and democratic orientations) a fact that gives further support to the meaningfulness of response consistency (cf. Lind, Sandberger & Bargel 1980).

Second we must face the fact that within the personality structure paradigm some traditionally computed measures of scale quality (e.g., reliability and validity) have to be interpreted in a new way. The concept of reliability has implications which are clearly at odds with the structural concept of personality. Since we can easily select individuals with known consistency we are in the position now to manipulate the “reliability” of the scale deliberately. (Table 3 gives two examples for this). In other words, the supposed characteristic of the scale is dependent on the kind of sample of persons we select. The empirical findings depicted in Table 3 also give support to the third, developmental, hypothesis. They show that the degree of consistency varies according to the stage of judgment to which the arguments belong. Since these stages are not defined by post hoc analysis but are pre-established on the basis of Kohlberg’s cognitive theory of moral development we have good reasons to conceive of consistency as a developmental variable. As outlined in another paper (Lind 1978a), consistency with regard to factor Stage increases with age and educational experience.

This finding of our analysis also sheds some light on the bewildering results of recent impact-of-college research (cf. Feldman & Newcomb, 1969). Some of this bewilderment may arise because most of these investigations have not distinguished between the affective and the structural component of attitude structure and thus have gained confounded results. Some of these investigations as well as our preliminary results suggest that, while university education may indeed leave the affective component largely unaltered, we may have to expect that higher education markedly contributes to the organization and reorganization of the cognitive structure of the students’ attitude system. The FORM-project provides a good basis for exploring this thesis and its implications with regard to a multitude of “questionnaire-domains”, e.g. moral judgment, socio-political values, perception and evaluation of environment, sense of competency, intolerance of ambiguity.

Anmerkungen

* Paper presented at the 8th International Meeting of the FORM-project in Portoroz, March 28-30, 1980.

(Broken table2)

Group 1 Determination by Factor STAGE: H 10% (N = 181) Stage Story Mercy Killing Story Theft

Internal Consistency^b Pro Con Pro Con I -.15^c-.03 .08 .09 -.11 II -.39 -.38 -.35 -.45 -.1.59 III -.24 -.15 -.22 -.02 -.43 IV -.41 -.38 -.41 -.30 -.1.48 V -.34 -.46 -.19 -.26 -.1.13 VI -.54 -.57 -.46 -.40 -.2.44

Tab. 3a: Item-Scale Correlations and Internal Consistency of Six-Stage-Scales, Applied to two Groups with different degrees of Stage-Structuring^a

Group 2 Determination by Factor STAGE: I 40% (N = 167)

Stage Story Mercy Killing Story Theft Internal Consistency^b Pro Con Pro Con I .41 .41 .21 .33 .53 II .34 .25 .23 .26 .45 III .47 .50 .46 .31 .64 IV .09 .14 .08 .17 .24 V .04 .03 .03 .07 .06 VI .04 -.03 .02 -.07 -.03

Tab. 3b: Item-Scale Correlations and Internal Consistency of Six-Stage-Scales; Applied to two Groups with different degrees of Stage-Structuring^a This, in comparison to earlier published data (e.g. Lind 1977), more conspicuous of coefficients results partly from the correction of a program mistake, partly from taking groups with greater contrasts with regard to their range of the factor Stage variable. b We took the deHoydt coefficient. c Negative coefficients are not interpretable.

Kapitel 4

EXPERIMENTAL QUESTIONNAIRES:
A NEW APPROACH TO PERSONALITY RESEARCH**1. Introduction*

The concept of Experimental Questionnaires is an attempt to facilitate new, adequate means for assessing dynamic-structural traits. In this paper I shall give, as briefly as possible, an account of the rationale, the design and the utility of the concept.¹ Traits and dispositions of individual persons are the genuine object of personality research. Contrary to Differential Psychology which embarks upon the assessment of any difference between persons, for personality research the assumption of behavior organizing traits is constitutive (Mischel 1977, p. 248). But, in spite of the simplicity of this insight, it remains a rather difficult task to translate the notion of traits into an unanimously accepted scientific concept. Hitherto, the empirical evidence pertaining to this issue has been ambiguous and has served both, those who believe in the reality of traits and those who deny it. While in some fields of personality research (test-)behavior is strikingly consistent, consistency is disappointingly low or absent in others. How can this persisting crisis be overcome? How can we cope with the stagnation in personality research? For a considerable time we believed, maybe with a few exceptions, that the crisis was due to a lack of either theoretical or methodological sophistication. We have witnessed a flood of new theoretical speculations and of new research strategies that are triggered by the failure of personality psychology to account for its perplexing findings. Doubtlessly, these endeavors have enriched personality research, yet we must admit that the "slow progress of soft psychology" (Meehl 1978) is still among the few facts on which psychologists would most easily agree. If the stagnation of personality research cannot be overcome by the improvement of either theory or method, the problem may be more fundamental. In fact it may be located in the relation between theory and method. As early as 1951, Travers pointed to the importance of theory guided construction of personality tests for the enhancement of psychological knowledge. Recently, Fiske (1970) mentioned again the "problem of coordinating the measure and the concept" (p. 51). The programs of "construct validation" (Cronbach & Meehl 1955) and of "substantial validation" (Loevinger 1957) have also this pro-

blem in mind, but unfortunately reduce it to the problem of optimizing the correlation between ‘similar’ measures. The cleavage between theory and method in personality research has led to two major ‘solutions’. Some eminent researchers have turned completely away from psychometrics and formed methods of their own, the so-called clinical methods (e.g. Freud, Piaget). The clinical interview method, if theory-related, doubtlessly deserves to be considered further on. Yet its lack of transparency and testability forces us, in addition, to search for more “objective” methods. (The expectational method of G.A. Kelly is somewhere in the middle and would need to be treated in a larger paper.) Most others apparently have adopted a dualistic view that conceives of theory and method as two somehow related yet not strictly connectable realms of research. This way, it is believed, one could escape the reproach of positivism and still take advantage of “objective” methods.

Allport (1951), one of the prominent advocates of this approach, admits however, that as a consequence “we must rest content with mere approximations to the structure of personality” (p. 422). So the latter approach is not satisfactory either. As Block (1968) sharply stated, “if a science of personality is to be formed, the responsibility of coupling concept and measure must be met” (30). For that we cannot borrow methods from elsewhere; “the task itself must create the method that is appropriate for it!” (Dilthey) Yet the problem of coordinating theory and method, in fact, is difficulty to achieve particularly in the field of personality research. Koffka (1934): “nowhere is it easier to miss the point, and run either into the scylla of blind statistical investigations of trait, or the Charybdis of ultimately unscientific abstract discussions” (677). Moreover, the solution of this problem cannot be confined to the construction of a ‘valid’ and ‘reliable’ assessment method but must also comprise a re thinking of the methodological criteria for test construction. We must be aware of the fact that concepts like ‘validity’ and ‘reliability’ may be inadequate themselves. Validity: Given that research methods are not ends in themselves but are always means for executing and testing a theoretical proposition, the Validity problem is twofold: (a) How empirically valid is the theoretical proposition? And (b) how theoretically valid (adequate) is the assessment method? In 1957 Loewinger pointed out the necessity of making this distinction, but unfortunately did not recognize the profundity of its implications for her cause and for psychometric methods in general. Such an epistemological point of view challenges the meaningfulness of the current practice of speaking of the empirical validity of tests’. Accordingly a test cannot be empirically validated, at least not directly. One may conclude that if a test has

been successfully employed for validating a theory it must have possessed theoretical validity. Note, however, that this conclusion is a-symmetric: A failure to produce supporting findings may be either due to a lack of empirical validity of the theory or to a lack of theoretical validity of the method or to both. Hence theories are falsified by empirical findings if and only if the method was theoretically valid (Lakatos’ objection to crude falsificationism). So, in order to prevent confusion, I rather prefer to speak of the test’s utility than of its (empirical) validity, when the test helped to produce ‘good’ results. Reliability: Lumsden (1974) has made a strong case for regarding the reliability problem also as a problem of a test’s validity. In fact the necessity of doing so is obvious since (a) classical test theory has advanced a substantial theory of inconsistencies of test behavior, namely error theory, and (b) this theory is testable, for it entails the assumption that this error is solely due to the test (and not to the properties of the individuals tested). When this objection is disregarded psychology runs the risk of carrying along hidden “anthropological assumptions” (Hilke 1980) which are at odds with the very theory involved. So

“the study of measurement must begin within discussion and analysis of such matters as the nature of traits and the mode of representation of the total personality” (Cattell 1946, p. 4).

2. Problems of translating personality theory into an adequate research methodology (Theory of trait assessment)

Experimental Questionnaires (EQs) are designed to fit the intentions of dynamic- structural theories of personality (e.g. Warren & Carmichael 1930; Lewin 1961; Allport 1961; Kohlberg 1969; but also many of the theoretical propositions of Stern 1911; Hartshorne & May 1928; Cattell 1946). According to these theories, Personality is the dynamic organization of the individual’s mind that determines (or better: that is) his or her characteristic interaction with the social environment at a particular level of development. This is not the place to discuss in depth the problems that result from translating this definition of personality into an adequate research methodology. But a condensed account of ten recurring problems of personality research may help to understand the rationale of the concept of Experimental Questionnaires and some of its implications. As far as I can see, these problems are, by and large, those which other researchers

in this field also encounter frequently: (1) the problem of conceptualizing traits, (2) the problem of inferring traits from behavioral acts, (3) the problem of functional ambiguity of acts, (4) the problem of separating the dynamic and structural components of a trait (5) the problem of conceptualizing person-situation interaction, (6) the problem of assessing meaningful behavior, (7) the problem of determining the experimental unit, (8) the problem of defining behavioral consistency, (9) the problem of locating the source of inconsistency, and (10) the problem of representativity. For the purpose of this paper these problems need not, and cannot, be extensively treated. It will suffice to make explicit the set of tentative solutions which constitute the rationale of EQs:

(1) Graumann (1960) juxtaposed two scientific concepts of trait: One is the “statistical”, which “is nothing else but a labeling of any frequent coincidence of particular, consistent, (behavioral) signs for the sake of better communication”. The other is the “hypothetical construct” concept of trait, which comprises only those correlations among signs that can be interpreted through recourse to a hypothesized common instance, i.e. through reference to a substructure, as meaningful associations” (147; translation mine). Resorting to a dynamic (teleological, organismic) definition of personality brings us, of course, closer to the latter, Internal-Structure (Allport), rather than to the former, Positivist, concept of a trait. Yet, since the latter definition seems to contain a moment of unwarranted speculation we prefer not to distinguish between ‘real’ behavior and ‘hypothetical’ traits. Surely, traits are different and need to be assessed in a different way. But this does not necessarily entail that they are less real, or more hypothetical, than acts (cf. Loevinger’s 1957, p. 642, critique of Cronbach and Meehl’s understanding of “hypothetical constructs”; also Hempel’s, 1958, discussion of “theoretical entities”). According to our conception, traits are not an instance ‘behind’ the behavioral acts, nor are they a ‘cause’ of acts but must be regarded as the combined whole of (a) dynamic goals (values, orientations), (b) instrumental acts, and (c) the relational structure between acts and goals (Lind, Sandberger & Bargel, 1980). We conceive of the formal ground-structure of dynamic-structural traits as being a little more complex than Miller et al.’s (1960) T-O-T-E model in so far as we regard the multiple determination of acts as being of great importance for theory (e.g. for conceptualizing ‘conflict’) and for research into personality (see below). But it is still rather simple so that the difference to other trait concepts can be seen easily (cf. the following graph):

(2) This model has a direct bearing upon the problem of inferring traits from behavioral acts. First, the relation between acts and dispositions should be conceived of neither as unknown, nor as stochastic nor as mediated by an intervening variable but as being of logical inclusion: Acts are constitutive parts of a trait. Second, the relation of acts to the two components of a trait, the dynamic and the structural, is well defined and can be reconstructed empirically if that particular trait really exists.

(3) Viewing acts as being multiply determined implies that when interpreting a single act we have to cope with its functional ambiguity (Kempf 1978, p. 12; Nunner-Winkler 1978, p. 352). Since there is no one-to-one relationship the problem of making inferences about a trait from a single observation is a substantial one. For example (cf. Lind 1978a), a person may agree to an argument of mine in defense of someone's decision - because it fits nicely into his opinion about the matter which he, or she, already held before (opinion agreement) - because the person feels, the argument is referring to a good reason (quality or level of reasoning) - or because it is me who has forwarded the argument (acquiescence) The index in parentheses alluded to the diversity of traits that may be involved, either alone or in combination, in producing a single behavior and that is of so great annoyance to the psychometrician who attempts to design a uni-dimensional test. This problem seems to be ubiquitous, as it is encountered in the research into moral attitudes (Nunner-Winkler 1978) as well as in intelligence testing (Kempf 1978; Loevinger 1957, p. 647) [see note 1].

(4) A problem pertains to the conceptual relation between the dynamic (valuing, motivational) component and the structural (cognitive) component of a trait, i.e. between the *Richtungs-* and *Rüstungs* component (direction and support component, W. Stern) of a disposition. According to the concept adopted both components are clearly distinguishable but not separable, We, therefore, do not speak of different 'dispositions' or 'faculties of mind' but of different components of one and the same trait. The theorem of inseparability entails also that neither component can be defined without reference to the other. While purely formal relations are empty, unrelated acts and values remain meaningless. This problem is, in my opinion, of great relevance for the prevailing discussions about (a) the 'content-specificity' of cognitive structure (complexity, style, etc.) and (b) the 'cognitive turn' of motivation and attitude research. In both instances, the superficial separation of the two components of traits apparently has resulted in severe difficulties with regard to conceptualization and measurement of personality variables.

(5) Another perplexing conceptual problem arises from the notion of person-situation interaction. In fact it seems as if we are even moving away from some promising attempts of early personality research. Though many refer today to the well-known formula of Lewin: $V = f(P \& U)$ (behavior is a function of person and environment interaction) only a few realize that this formula does not imply that, as 'modern' interactionism surmises, person and environment can be empirically separated. Lewin (1961) did not mean that we should conceive of separate entities but emphasized their interaction, i.e. 'life-space', as being the object of psychological research. Accordingly, personality is the way a person perceives, evaluates, and reacts to a situation (Helm 1960). Hence "in concrete experiments both cannot be separated" (Olweus 1976, 65). Maybe, that our language is responsible for some of the prevailing confusion as it does not distinguish between a separating and a combining meaning of the word 'and' [2].

(6) Personality research is not concerned with the mere physical properties of human behavior – physiology and other branches of science take care of this – but with meaningful behavior. Verbal behavior is not the only kind of meaningful behavior but it may reasonably be considered its most preferred medium. Verbal behavior, however, alone does not suffice to represent meaning. In addition we need to ascertain its functional and structural properties with regard to the 'totality' of a trait.

(7) Nothing seems to be more obvious than that the unit of personality research is the individual. In theory, we almost exclusively refer to the individual personality. There seems to be no meaningful way of setting the 'personality of a sample of persons' equal to the individual personality structure. In practice, however, it is often implicitly or explicitly assumed that traits are commonly organized. Since such a notion does not take into account the individuality of traits, Allport (1961) has aptly called this the External-Effect definition of traits. Persons are judged merely according to some external, socially defined standards of conduct and achievement. This is an approach in its own right but it does not accord with our concept of personality.

(8) In many text-books one can find the statement that personality development means both integration and differentiation. In empirical research this often is translated as: Behavior becomes more and also less consistent. An obvious contradiction! In fact research has produced supporting evidence for both propositions. A closer look shows that at the heart of this problem there is an insufficient definition of 'consistency'. If 'consistency' means "holding to the same principle or

practice” (Webster’s New World Dictionary, a proposition which does not specify the ‘principle’ or ‘practice’ is incomplete. In this context ‘consistency’ is a relational concept. A behavior cannot be consistent per se, but only ‘consistent with regard to something’; in our case ‘with regard to the function it is to serve’. So behavior may become more consistent with regard to function A and, at the same time, become less consistent with regard to function B.

Besides such simple processes of integration and disintegration there may also be a superposition of traits, whereby the superposed trait does not become disintegrated but ‘differentiated’ with regard to another ‘qualifying’ trait. The problem of defining consistency also has a bearing upon the distinction between the External-Effect and the Internal-Structure model of trait (Allport). While the former assumes that behavior must be, if at all, consistent with regard to some external standard, the latter assumes the reality of internal standards to which behavior may, or may not, accord.

(9) One way to save the External Effect (common trait) model in the face of an overwhelming number of discrepant facts is to assume that the measurement process introduces random error, i.e. that inconsistency of behavior with regard to a hypothesized common trait is due to a property of the assessment method. This assumption, advanced by classical psychometrics (Spearman, Gulliksen, Guilford etc.), raises the problem of attributing inconsistency. While ‘error theory’ attributes inconsistency to the ‘reliability’ of the method, the theory of dynamic-structural personality conceives of it as an attribute of trait organization. This is an empirically testable question! In fact, research into personality and personality development has produced ample evidence that inconsistency of behavior varies systematically with the kind of trait measured and with the phase or stage of personality development (see, for example, Hartshorne & May 1928, Kohlberg 1969, Loevinger 1976, Lind 1978a, [Lind 2019]). In sum, I think, we should not rule out the assumption of random error (unreliability). This precisely defined concept (Kempf 1978), which is to be distinguished clearly from the trivial explanation of ‘unknown factors’, may, however, come in as the last, and not as the first, explanation for observed inconsistencies. (10) Finally we need to provide a tentative solution for the problem of representativity. Brunswik (1955), as many psychologists after him, has invoked the concept of ecological validity which means that the experimental stimuli should be distributively representative. But what population of stimuli should we refer to when designing an experiment? There are uncountable possibilities, some of which are unassessable in principle (e.g., the distribution of events over time). So how to

chose? I believe that this problem is spurious with respect to theory-guided research. Generalizations can never be based merely on a single experiment but are always based on the comprehensive experience comprised in a scientific theory. In order to test the empirical validity of a hypothesis we need rather to ascertain categorial representativity. Only this guarantees that the research instrument tests what it is supposed to test. I, therefore, believe that many objections against systematic experimental design are truly directed at the positivist spirit in which they are often used. Like Lewin (1963), I think that “experiments become superficial only if just one or the other condition is operationalized but not the essential structure” (p. 200). What is “essential” must, of course, be determined by a theory. It is noteworthy that Brunswik (1955) has not only introduced the concept of ‘ecological validity’ but also shown very convincing causes for the epistemological utility of systematic research designs (the “diacritical method”). For the choice of a theoretically valid research strategy our previous assertions have three basic implications. First we must conceive of the research situation as being part of the personality investigated. Adequate research methods need not eschew interaction with the subjects but may rather profitably make use of those individual and concrete person-environment constellations which conspicuously expose their dynamic characteristics” (Helm 1960, p. 374).

Second, according to the dynamic-structural trait concept “the task of psychometrics is to isolate, to identify and, so far as possible, to measure separately (though not as separate objects!) the important components of (behavioral) variance” (Loevinger 1957, p. 649; parentheses added). We would add: The task is to translate a particular personality theory into an adequate research situation which allows the researcher to infer the structural and the dynamic component of the hypothesized traits, i.e. their behavioral relevance (‘cognitive anchoring’) and their directional value (attitude, orientation). Third, the problem of multiple determination of acts, and the resulting functional ambiguity of acts, should not be viewed as being insolvable in principle (‘indeterminism’) but can be coped with through applying the ‘diacritical method’ (Brunswik). I believe this hermeneutic device is well suited to sort out to which of several functional unities (traits) an act belongs. It requires that a pattern of responses to a systematically (orthogonally) designed pattern of probing questions is analyzed as to which trait, or set of traits, account for that particular person-situation interaction. With regard to these three points, I think, the question of whether trait assessment is executed by a skillful clinician through personal interview or by a

survey researcher through carefully prepared questionnaires is of minor importance (Lind & Wakenhut 1980).

The variation of theoretical validity within each approach seems to be greater than between them. There are, however, some good reasons why personality research cannot rest solely on highly elaborate and extensive clinical interviews. The lack of sufficient transparency and systematics, which makes it all too easy for uninformative circularities and trivial explanations to creep in, is not the least important reason for attempting a new, complementary, approach to personality research.

3. The design of Experimental Questionnaires

At the first sight this method appears to be a contradiction in itself. Experimental Questionnaires (EQs) neither fit into usual methodological categories, e.g. Cattell’s distinction of L-, Q-, and T-

techniques, nor are they a new way of utilizing questionnaire methods within experiments. This method rather combines, maybe in a surprising way, the meaningfulness and economy of the questionnaire technique with the epistemological advantages of systematic, multifactorial design. Unlike self-report questionnaires, EQs are conceptualized as “objective tests” in Loewinger’s (1957) sense, namely in the sense “of structures tests viewed behavioristically” (648). The behaviorism of EQs, however, is not a Watsonian or a Skinnerian one but more like the “subjective behaviorism” of Miller et al. (1960) and the “social behaviorism” of G. H. Mead (1968/1934).

EQs are viewed “subjectively” because they are designed to penetrate through the surface of unrelated behavioral acts and to assess the dynamic trait structure of individuals. EQs are, nevertheless, also social in so far as their scientifically construed categories are bound to the sociality of communication. EQs are used nomothetically (i.e. they seek systematic regularities of human behavior), but, at the same time, they possess ideographic sensibility [3]. EQs are designed to answer two questions of measurement in this order: 1. To what degree does a hypothesized trait account for an observed pattern of responses of an individual? and 2. in which direction does this factor influence the test response? While the first question aims at the assessment of the cognitive-structural component of a particular trait, the second question aims at a trait’s dynamic-motivational component. In addition, EQs may be designed as a multi-factorial

experiment so that more than one dynamic trait can be incorporated in order to study a more complex system of traits. This way EQs also facilitate means to cope with the problem of functional ambiguity. Through multifactorially designed EQs it become possible to find out which of several traits functions alone, or in combination with other traits, as a frame of reference for a particular person-situation interaction. In principle, an unlimited number of traits can be analysed diacritically by EQs, although in practice the rule of parsimony (Ockham's razor) may be applied which entails starting always with as few hypothesized traits as possible. As an example for EQs, the three-factorial design of the ["Moral Competence Test" (MCT)] is depicted in Table 1. The aim of this test is to assess the cognitive and the evaluative component of an individual's moral consciousness.

The MCT has been developed on the basis of Kohlberg's theory of moral-cognitive development (Kohlberg 1969; Rest 1973; Lind 1978a, 1980e, [2019]), and on a theory of response behavior which contains three tentative assumptions derived from recent moral judgment research (Keasey 1974, Lind 1977, [2019]): These assumptions assert (1) that an individual's evaluation of moral arguments for and against a particular moral decision is determined by the quality of the moral argument, i.e. by the individual's attention to the Stage-type of moral reasoning exhibited by the arguments (according to Kohlberg 1969); (2) that the statements may be also evaluated with regard to their agreement or disagreement with one's own opinion about the moral dilemma; and (3) that inconsistency of judgment behavior may not just indicate a lack of moral development but a greater differentiation; it may indicate (a) the successful co-ordination of the moral consciousness factor and the opinion agreement factor ('mature moral commitment'; Perry 1970); and (b) the differentiation of the preferences for a particular level of moral reasoning according to the situation context in which the moral decision is made ('contextual relativism'). The test-version depicted in Table 1 contains two subtests (Theft and Mercy Killing). Each is introduced by a brief description of a dilemma of action and by a subsequent question as to how "wrong" or "right" the respondent judges the solution chosen by the protagonist. Thereafter six pro-arguments and six con-arguments are presented, each at random order. Thus the [MCT] forms a 2x2x6 ideographic experiment. For scoring the [MCT] we have surveyed several computational procedures (Lind et al. 1980, Lind 1980g).

Hitherto the analysis of variance components proved to be one of the most adequate and fruitful methods. It enables the psychologist to quantify the degree to which a certain hypothesized trait,

alone or in conjunction with others, determines an individual's pattern of response behavior (the structural-cognitive component). The evaluative component is scored as usual by summated ratings (average response).

4. Some concluding remarks on the empirical utility of Experimental Questionnaires

We have argued that prevailing standards of test evaluation do not apply to assessment methods like Experimental Questionnaires. But this does not mean that there are no other ways to ascertain the goodness of a research method. I can see at least two possibilities: One is concerned with the theoretical validity of the assessment technique. Since we have 'deduced' the concept of Experimental Questionnaires from a theoretical proposition this method may legitimately be evaluated with regard to that claim. The other criterion involves both theory and method. The choice of the dynamic-structural definition of traits entails that a test based on this definition is an empirically useful device for personality research in so far as it widens the 'categorical overlap' (Campbell 1963) between the researcher and his, or her, subjects. The greater this categorical overlap is, the better, we can say, we understand a particular person-situation interaction. I believe that both criteria provide good means for criticizing the method of Experimental Questionnaires, and thus foster the progress of personality research.

(Broken table)

Independent Variables (Factors) Dependent Variable 1. Story (Dilemma) 2. Stage of reasoning 3. Pro-Con (Opinion-Agreement) Judgement of Acceptability - A - I - II - III - IV - V - VI Pro³ Con ••••• compelety completely unacceptable acceptable -4 -3 -2 -1 0 1 2 3 4 - B - I - II - III - IV - V - VI Pro³ Con •••••

Tab. 1: Experimental Questionnaire: The design of the "Moral Competence Test" (MCT)

Theoretical validity, of course, cannot be quantified but must be established on the grounds of conceptual analysis and theoretical reflection. A profound explication of the respective theory as well as some kind of expert-rating of the resulting test are the major devices for securing theoretical validity. This has been done in the case of the [MCT]. So for the time being we shall assume that it is theoretical valid [4]. With respect to the categorial overlap of Experimental

Questionnaires there is some quantitative evidence available. The [MCT] which is based on the EQ- methodology has been employed recently in a number of studies comprising several hundred of subjects. One of the most important results is that some propositions of Kohlberg's (1969) theory of cognitive-moral development were rather well corroborated (Lind 1980d). A second important result is that it may suffice to take into account only a few dynamic-structural traits in order to understand a person's judgment behavior in a moral situation. This is of relevance for the utility criterion. Ideally the hypothesized trait-structure of moral consciousness should account completely for the variance of the individual response behavior in the experimental situation. This does not necessarily imply that behavior is completely predictable, but that we can comprehend it as meaningful given the knowledge of a particular trait structure. Past personality research has shown that such a complete overlap between theory and reality is far from being reached. Hartshorne and May's (1928) Studies in the Nature of Character revealed only minimal categorial overlap; the squared intercorrelations of children's behavior across different 'moral' situations, which was taken as an indicator for this, was as low as .05. This is even below the .10, which Mischel (1968) has identified as a barrier to personality research.

(Broken Table)

Approach (- Example) Categorial Overlap* r^2 1. Common Trait, External-Effect definition - Harthshorne & May (1928) - Mischel (1968) .05 .10
 2. Restricted Common Trait definition - Bem & Allen (1974) .20 3. Internal-Structure definition, dynamic-structural traits - Lind et al. (1980) .27
 * Squared correlational coefficient, i.e. proportion of variance of judgment accounted for by moral consciousness, or personality traits, respectively.

Tab. 2: 'Categorial overlap' in three approaches to personality research: a tentative comparison

G. Allport criticized this study for being a typical representative of the External- Effect definition of traits. Children may not behave consistently with regard to some socially defined criteria and still be consistent with regard to their own motives. Bem and Allen (1974) showed, though with regard to another topic, that consistency increases up to .20 when the individual's awareness of trait consistency is taken into account. Since EQs represent a new approach which claims greater adequacy with regard to the dynamic-structural trait concept we should expect a better agreement between the hypothetical model and the reality of an individual's reasoning in a moral situation. In fact, in a study of German high school graduates ($n = 516$) one trait (moral consciousness) already accounts on average (median) for .27 of the individuals' behavior variance. There is also

evidence that this number varies with the level of moral development; the less people are developed the less they apply moral dimensions in their judgment behavior (Keasey 1974). But this can also be looked at from another angle. Usually a trait is assessed, regardless of whether there is a large, medium or no categorial overlap between the researcher and the particular subject. If the consistency of test-behavior across a sample of persons is not perfect, according to most textbooks, this can be coped with by enlarging the number of items, unless reliability is zero. Besides the frequent experience that more items often add more 'heterogeneity' to the assessment device and thus, in practice, often do not improve it, this strategy is basically problematic as noted above. In many cases we 'measure' traits that do not exist in subjects and do not even notice it. Someone who does not 'have' the particular trait measured gets the same score as the person who exhibits a medium (or neutral) 'amount' of this trait. E.g., a person who is as much introvert as extrovert in his responses to individual items of a scale receives the same, average, score as a person who responds consistently medium extrovert, or medium introvert.

In assessing the degree to which a hypothesized dynamic component organizes instrumental acts into a functional whole, Experimental Questionnaires avoid this pitfall. This method provides a measure for ascertaining the degree of categorial overlap between an individual person and the researcher before the respective category is applied for classifying this person. The person is thus not forced into a descriptive system which actually does not tally with his, or her, way of perceiving the world. The measure of 'cognitive anchoring' of a trait thus gives us also an opportunity of becoming aware of the descriptive adequacy or inadequacy of the theory on which the assessment technique is based. It tells us something about the ability of a psychological theory to understand particular persons and persons in general. I am deliberately not speaking of the psychologist's ability to understand since this may deviate greatly from the common knowledge as codified in a theory. Certainly, and hopefully, many psychologists also make successful use of other sources of knowledge.

Notes

[1] There are similar problems in natural sciences. As Piaget (1973) observes, also in physics, chemistry and biology “at the outset the problem is always to isolate the components out of the bewildering mass of phenomena” (36; translation mine). See also Popper (1979).

[2] Again this problem is not unique for psychology. Also in the natural sciences the “Observable is only the result of an interaction between experiment and reality” (Piaget 1973, 52). Hence the scientist’s major task is to disentangle the nature of the observation i.e. the observation mediating processes) and the nature of the observed. This is, even in physics, less trivial than might be assumed. Heisenberg, for example, interpreted his indeterminacy principally for a long time as pertaining to the measurement process while it was only eventually realized that it pertained to the nature of a quantum (cf. Popper 1979, 305).

[3] Unfortunately ‘nomothetic science’ is sometimes misunderstood as postulating a priori the lawfulness of any behavior. This is, of course, not the case. Nomothetic sciences are rather law-seeking. They critically test the question to which extent the hypothesis of lawfulness is empirically valid. This implies also that sciences have to be aware of idiosyncracies.

[4] Note that EQs always require the existence of an explicit substantial theory (which is more than a statistical account of singular investigations or a mathematical model). Otherwise there is no reason for employing EQs. On the other hand a theoretically valid EQ can be regarded as the explication of a theoretical concept and, therefore, constitute the meaning of the concept. If there was no concept from the outset, such an ‘operational’ definition would be meaningless.

Anmerkungen

* Paper presented at the 22nd International Congress of Psychology in Leipzig (GDR), July 1980. (Revised and extended Version). 1) I like to thank a number of colleagues inside and outside of our research group for their co-operativeness and their sincere encouragements, in particular Wilhelm Kempf (U o Konstanz), Marc Lieberman (Harvard U), Paul Meehl (U o Minnesota), William Perry (Harvard U) and Roland Wakenhut (Sozialwissenschaftliches Institut 50

der Bundeswehr). Their comments on earlier versions of this paper and their willingness to discuss special topics in detail have very much helped to sharpen the ideas advanced therein. For remaining soft spots, however, I take sole responsibility. [Institutions added, 2019]

Note [2019]

Even after almost 30 years, I feel no need to change anything which I have written here. I only would do some terminological changes. The “MUT” or “m-u-t” (Moralische Urteils-Test; moral judgment test) has in meanwhile be renamed to *Moral Competence Test* (MCT). I changed this already in the text. Also the word “component”, which I took from literature, would not go without comment and correction today. Components are things, objects, bodies and alike. Yet, in

psychology we actually deal with attributes or aspects of behavior, not with components. As I argue in later writings, this substantiation of aspects has misled many generations of psychologists and social scientists to look for the wrong “thing,” and to design wrong instruments of measurement (Lind 2016).

Literatur [ergänzt 2019]

- Adelson, J. & O'Neill, R.P. (1966) Growth of political ideas. The sense of community. *Journal of Personality and Social Psychology* 4, 295-306.
- Allport, G. (1961) *Pattern and growth in personality*. New York: Holt.
- Bargel, T., Framhein, G., Kellermann, P., Peisert, H. & Sandberger, J.-U. (1973) Zweckbestimmung des Studiums als Thema der Hochschulforschung. *Zeitschrift für Pädagogik* 6, 932-965.
- Bargel, T. & Framhein, G. (1976) Zur Diskussion von Bildungszielen und zur Leistungsmessung im Hochschulbereich. In: W. Zapf (Hrsg.), *Gesellschaftspolitische Zielsysteme. Soziale Indikatoren IV*. Frankfurt: Campus, 126-161.
- Becker, H., Geer, B., Hughes, E. & Strauss, A. (1961) *Boys in White: Students culture in medical school*. Chicago: University of Chicago Press.
- Bem, D.J. & Allen, A. (1974) On predicting some of the people some of the time. *Psychological Review* 81, 506-520.
- Block, J., (1968) Personality measurement. Overview. In: *International Encyclopedia of the Social Sciences*, Vol. 12, 30-36.
- Brennan, J. & Franke-Wikberg, S. (1980) The socialization outcomes of higher education. Some problems of methodology. Paper presented at the Sociology of Education Research Committee of the Int. Sociological Association, Paris, August 1980.
- Bromberek, B. (1978) *Die Untersuchungskonzeption: Die Sozialisation der Studenten in einer Hochschule - ihr Verlauf und ihre Ergebnisse*. Vienna: European Coordination Center, FORM-Projekt, Document 17 (revised).
- Broughton, J.M. (1978) Criticism of the developmental approach to morality. A.P.A. *Journal Supplement Abstract Service*, November 1978.
- Brügelmann, H. (1979) Erfahrungen mit dem Test zum "Moralischen Urteil" (M-U-T) bei Jura-Studenten. In: G. Lind (Hg.), *Moralische Entwicklung und soziale Umwelt*. Tagungsbericht. Universität Konstanz. [The M-U-T has been renamed as Moral Competence Test, MCT]
- Brunswik, E. (1955) Representative design and probabilistic theory in a functional psychology. *Psychological Review*, 62, 193-217.
- Campbell, D.T. (1963) Social attitudes and other acquired behavioral dispositions. In: S. Koch (Ed.), *Psychology: A study of a science* New York: McGraw-Hill, 94-172.
- Cattell, R.B. (1946) *The description and measurement of personality*. New York: World Book.
- Cattepoel, J. (1977) *Sozialstruktur und Verhalten der Studienanfänger der Jahrgänge 1975 und 1976*. Erster Zwischenbericht FB Rechtswissenschaft, Universität Trier.
- Converse, P.E. (1970) Attitudes and Non-Attitudes: Continuation of a dialogue. In: E.R. Tuftte (Ed.), *The quantitative analysis of social problems*. Reading: Addison-Wesley.
- Cronbach, L. & Meehl, P. (1955) Construct validity in psychological tests. *Psychological Bulletin* 52, 281-301.
- Durkheim, E. (1976) *Soziologie und Philosophie*. Frankfurt: Suhrkamp (Original 1924).
- Feldman, K. A. & Newcomb, T. M. (1969) *The impact of college on students*. Vol. I. & II. San Francisco: Jossey-Bass.
- Fiske, D. & Pearson, P. H. (1970) *Theory and Techniques of Personality Measurement*. *Annual Review*, 49-86.
- Forschungsgruppe *Hochschulsozialisation* (1979) *Wissenschaftlicher Bericht 1976-1978*. Zentrum I Bildungsforschung, Universität Konstanz.

- Framhein, G. (1980) Report of a Special Seminar on University students: their training and conception of life. FORM-papers presented at the 3rd EARDHE Congress at Klagenfurt, January 1979. Forschungsgruppe *Hochschulsozialisation*. Universität Konstanz.
- Graumann, C. F. (1960) Eigenschaften als Problem der Persönlichkeitsforschung. In: Ph. Lersch & H. Thomae (Hg.), *Handbuch der Psychologie*, Bd. 4, Persönlichkeitsforschung und Persönlichkeitstheorie, Göttingen: Hogrefe, 87-154.
- Harten, H. Chr. (1977) *Kognitive Sozialisation und politische Erkenntnis*. Weinheim: Beltz.
- Hartmann, H. & Wakenhut, R. (1973) Über Aktualität und Effizienz von Attitudenskalen. *Zeitschrift für Sozialpsychologie* 4, 195-206.
- Hartshorne, H. & May, M.A. (1928) *Studies in the nature of character*. Volume 1. *Studies in deceit*. New York: Macmillan.
- Helm, J. (1960) Über Gestalttheorie und Persönlichkeitspsychologie. In: Ph. Lersch & H. Thomae (Hg.), *Handbuch der Psychologie*, Band 4, Persönlichkeitsforschung und Persönlichkeitstheorie, Göttingen: Hogrefe.
- Hempel, C.G. (1958) The theoretician's dilemma: A study in the logic of theory construction. In: H. Feigl, M. Scriven & G. Maxwell (eds.), *Minnesota Studies in the philosophy of science*. Vol. II. Minneapolis: University of Minnesota Press, 37-98.
- Hess, R.D. & Torney, J.V. (1967) *The development of political attitude in children*. Chicago: Aldine.
- Hilke, R. (1980) Zu den impliziten "anthropologischen" Voraussetzungen der klassischen Testtheorie. *Diagnostica* XXVI, 99-111.
- Höffe, O. (1975) *Strategien der Humanität. Zur Ethik öffentlicher Entscheidungsprozesse*. Freiburg: Alber.
- Hogan, R., DeSoto, C.B. & Solano, C. (1977) Traits, tests, and personality research. *American Psychologist* 32, 255-264.
- Huber, L. (1975) Anlässe, Ziele und Verlauf der Tagung 'Sozialisation in der Hochschule'. In: T. Bargel, G. Framhein, L. Huber & G. Portele (Hg.), *Sozialisation in der Hochschule*. Reihe *Blickpunkt Hochschuldidaktik* 37, Hamburg, S. 1-18.
- Ijzendoorn, M.H.v. (1979) Moralität, Kognition und politisches Bewusstsein. 54 *Zeitschrift für Pädagogik* 25, 547-567.
- Jacob, Ph. (1957), *Changing Values in College*. New York: Harper & Row.
- Keasey, Ch.-B. (1974) The influence of opinion-agreement and qualitativ supportive reasoning in the evaluation of moral judgments. *Journal of Personality & Social Psychology* 30, 477-482.
- Kempf, W. (1978) Kritische Bemerkungen zu impliziten Voraussetzungen psychologischer Testtheorie und -praxis und ihre Angemessenheit zur Erfüllung der Aufgaben curricularer Evaluation. Vortrag zum Symposium "Aspekte von Lehrplanevaluation", Veldhoven (Holland).
- Kohlberg, L. (1969) Stage and sequence. The cognitive-developmental approach to socialization. In: D. Goslin (Ed.), *Handbook of socialization theory and research*. Chicago: Rand McNally. (dt.: ders: *Zur kognitiven Entwicklung des Kindes*. Frankfurt: Suhrkamp, 1974.)
- Kohlberg, L. (1979) The meaning and measurement of moral development. Heinz Werner Memorial Lecture. Manuscript. Harvard University, Cambridge/Massachusetts.
- Kohlberg, L., Kauffman, Scharf & Hickey (1974) *The Just Community Approach to Corrections: A Manual, Part II*, Manuskript. Harvard, Cambridge.
- Krämer-Badoni, T. & Wakenhut (1978) Möglichkeiten der Skalierung des moralischen Urteilsverhaltens. In: L. Eckensberger (Hrsg.), *Entwicklung des moralischen Urteilens – Theorie, Methode, Praxis*. Universität des Saarlandes, Universitätsdruck, S. 379-391.
- Krausser, P. (1968) *Kritik der endlichen Vernunft. Diltheys Revolution der allgemeinen Wissenschafts- und Handlungstheorie*. Frankfurt: Suhrkamp.

- Kurtines, W. & Greif, E.B. (1974) The Development of moral thought: Review and Evaluation of Kohlberg's approach. *Psychological Bulletin*, 81, 453-470.
- Lewin, K. (1963) *Feldtheorie in den Sozialwissenschaften*. Bern/Stuttgart: Huber (Original: *Field theory in social sciences*. New York: Harper 1951).
- Lind, G. (1977) Über die Grundlagen einer integrationistischen Meßtheorie zur Erforschung des moralischen Urteils. Arbeitsunterlage 41, Forschungsgruppe *Hochschulsozialisation*. Universität Konstanz.
- Lind, G. (1978a) Wie mißt man moralisches Urteil? Probleme und alternative Möglichkeiten der Messung eines komplexen Konstrukts. In: G. Portele (Hg.), *Sozialisation und Moral*. Weinheim: Beltz, S. 171-201.
- Lind, G. (1978b) Some considerations on change, learning and development in higher education – theoretical and methodological issues related to the FORM-Projekt. FORM-Doc. 35
- Lind, G. (1980a) Zum Stand der moral judgment Forschung – Überlegungen 55 zum Gehalt und zur Bewährung von Kohlbergs Theorie. In: R. Silbereisen (Hg.), *Newsletter Soziale Kognition* 3. Berlin: Technische Universität, S. 75-85.
- Lind, G. (1980b) Reliability of test of structure of personality? Empirical findings on the problem of interpreting judgment consistency. (In diesem Band).
- Lind, G. (1980c) Experimental Questionnaires: A new approach to personality research. In: A. Kossakowski (Ed.) *Progress in personality research* (In diesem Band). (Vgl. Veröffentlichungsnachweis).
- Lind, G. (1980d) Toward an evaluation of the impact of university on students' moral development. In: E. v. Trotsenburg (Ed.), *Higher Education – A Field of Study*, Vol. 5. Frankfurt: Lang. (In press)
- Lind, G. (1980e) Moralisches-Urteil-Test (m-u-t), Hinweise zum Einsatz und zur Interpretation eines Forschungsinstruments. Arbeitsunterlage 68, Forschungsgruppe *Hochschulsozialisation*: Universität Konstanz.
- Lind, G. (1980f) Zur Bestimmung des Entwicklungsstandes der moralischen Urteilskompetenz beim Übergang vom Gymnasium auf die Universität. In: J. Domnick (Hg.), *Aspekte grundlagenorientierter Bildungsforschung*. Forschungsberichte 18, Zentrum I Bildungsforschung, Universität Konstanz, S. 151-166.
- Lind, G. (1980g) Intra-Individuelle Varianzkomponentenzerlegung der Antwortmuster des "Moralisches-Urteil-Tests" (Algorithmus für das Programmsystem KOSTAS). Arbeitsunterlage 65, Forschungsgruppe *Hochschulsozialisation*, Universität Konstanz.
- Lind, G. (2002). *Ist Moral lehrbar?* Berlin: Logos.
- Lind, G. (2016). *How to teach morality. Promoting deliberation and discussion, reducing violence and deceit*. Berlin: Logos.
- Lind, G. (2019). *Moral ist lehrbar. Wie man moralisch-demokratische Kompetenz fördern und Gewalt und Betrug verringern kann*. Berlin: Logos.
- Lind, G., Bargel, T. & Sandberger, J.-U. (1976) Die Vermittlung von über und außerfachlichen Qualifikationen während des Studiums. Arbeitsunterlage 28, Forschungsgruppe *Hochschulsozialisation*. Universität Konstanz.
- Lind, G., Nielsen, A. & Schmidt, U. (1976) Moralisches Urteil und *Hochschulsozialisation* - Materialien, Beiträge. Arbeitsunterlage 40, Forschungsgruppe *Hochschulsozialisation*. Universität Konstanz.
- Lind, G., Sandberger, MU. & Bargel, T. (1980) Moral judgment, egostrength and democratic orientations – Some theoretical contiguities and empirical findings. Arbeitsunterlage 64, Forschungsgruppe *Hochschulsozialisation*. Konstanz.(Vgl. Nachtrag).
- Lind, G., Sandberger, J.-U. & Bargel, T. (1983) Moralisches Urteilen, Ich- Stärke und demokratische Orientierungen. In: G. Lind, M.A. Hartmann & R. Wakenhut (Hg.); *Moralisches*

- Urteilen und soziale Umwelt. Weinheim: Beltz, S. 43-58.
- Lind, G. & Wakenhut, R., 1980: Erfassung von moralischem Urteil mit standardisierten Fragebogen. *Diagnostica* XXVI, 312-334.
- Loevinger, J. (1957) Objective tests as instruments of psychological theory. *Psychological Reports* 9, 635-694.
- Loevinger, J. (1976) *Ego Development*. San Francisco, Jossey-Bass.
- Lumsden, J. (1976) Test theory. *Annual Review of Psychology* 27, 251-280.
- Lippert, E. & Wakenhut, R. (1978) Zur Zentralität von Einstellungen. *Zeitschrift für Soziologie* 7, 87-96.
- Marton, F. (1978) Describing conceptions of the world around us. Reports from the Institute of Education, University of Göteborg, Sweden.
- Mead, G. H. (1968) *Geist, Identität und Gesellschaft*. Frankfurt: Suhrkamp (Originally published: 1934).
- Meehl, P. E. (1978) Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft Psychology. *Journal of consulting and clinical Psychology* 46, 806-834.
- Miller, G. A., Galanter, E. & Pribram, K. H. (1960) *Plans and the structure of behavior*. New York: Holt.
- Mischel, W. (1968) *Personality and assessment*. New York: Wiley.
- Mischel, W. (1977) On the future of personality measurement. *American Psychologist*, 246-254.
- Newcomb, Th. (1943) *Personality and social change: Attitude formation in a student community*. New York: Dryden Press.
- Nitsch, W. (1975) Zur theoretischen und politischen Diskussion um ein Forschungs- und Handlungsfeld Hochschulsozialisation. In: T. Bargel et al. (Hg.) *Sozialisation in der Hochschule*. Reihe Blickpunkt Hochschuldidaktik 37, Hamburg, 30-39.
- Nunner-Winkler, G. (1978) Probleme bei der Messung des moralischen Urteils mit standardisierten Verfahren. In: L. H. Eckensberger. *Entwicklung des moralischen Urteilens*. Universität des Saarlandes, Saarbrücken: Universitätsdruck, S. 337-358.
- Olweus, D. (1976) Der moderne Interaktionismus von Person und Situation und seine varianzanalytische Sackgasse. *Zeitschrift für Entwicklungspsychologie und pädagogische Psychologie* 8, 171-185.
- Oser, F. (1980) *Die Theorie von Lawrence Kohlberg im Kreuzfeuer der Kritik (Eine Verteidigung)*. Manuskript. Universität Fribourg, Schweiz.
- Piaget, J. (1973) *Das moralische Urteil beim Kinde*. Frankfurt: Suhrkamp (Original 1932).
- Piaget, J. (1973) *Erkenntnistheorie der Wissenschaften vom Menschen*. Frankfurt: Ullstein.
- Perry, W. G. (1970) *Forms of intellectual and ethical development in the college years*. New York: Holt.
- Phillips, D. C. & Nicolayev, J. (1978) Kohlbergian moral development. A progressing or degenerating research program? *Educational theory* 28, no. 4, 286-301.
- Popper, K. (1979) *Objective Knowledge. An evolutionary approach*. Oxford: At The Clarendon Press (Revised Edition).
- Portele, G. (1979) Soziale Vorstellungen von Wissenschaftlern verschiedener Disziplinen über Wissenschaft und Moral. In: G. Lind (Hg.), *Moralische Entwicklung und soziale Umwelt*. Tagungsbericht. Universität 57 Konstanz.
- Rawls, J. (1975) *Eine Theorie der Gerechtigkeit*. Frankfurt: Suhrkamp. Rest, J.R. (1973) The hierarchical pattern of moral judgment: a study of patterns of comprehension and preference of moral stages. *Journal of Personality* 41, 86-109.
- Reichhardt, C. & Cook, T. (1979) Beyond qualitative versus quantitative methods. In: T. Cook & C. Reichhardt (eds.), *Qualitative and quantitative methods in evaluation research*. London: Sage, 7-32.

- Reichwein, R. (1980) Identität und postsekundäre Sozialisation in soziologischer Sicht. Internes Arbeitspapier. Münster.
- Riesman, D. (1963) The influence of student culture and faculty values in the American College. In: H.M. Ruitenbeek (ed.), *Varieties of modern social theory*. New York: Dutton, 319-340.
- Sandberger, J.-U. (1979) Zu Struktur und Relevanz von soziopolitischen Grundwerten – Am Beispiel von Abiturienten. In: H. Klages & P. Kmiecik (Hg.) *Wertwandel und gesellschaftlicher Wandel*. Frankfurt, New York: Campus.
- Sandberger, J.-U. (1980) Aspekte des Wertsystems und des Gesellschaftsbildes von Hochqualifizierten im internationalen Vergleich. In: J. Domnick (Hg.), *Aspekte grundlagenorientierter Bildungsforschung*. Konstanz: Forschungsberichte 18, Zentrum I Bildungsforschung, Universität Konstanz.
- Sandberger, J.-U. & Lind, G. (1979) The outcomes of university education: Some empirical findings on aims and expectations in the Federal Republic of Germany. *Higher Education* 8, 179-203.
- Schluchter, W. (1979) *Die Entwicklung des okzidentalen Rationalismus. Eine Analyse von Max Webers Gesellschaftsgeschichte*. Tübingen: Mohr.
- Stern, W. (1911) *Die Differentielle Psychologie in ihren methodischen Grundlagen*. Leipzig: Barth.
- Thurstone, L.L. (1931) The measurement of social attitudes. *Journal of abnormal and social Psychology* 26, 249-269.
- Travers, R.M. (1951) Rational hypotheses in the construction of tests. *Educational and psychological Measurement* 11, 128-137.
- Warren, H.C. & Carmichael, L. (1930) *Elements of human psychology*. Boston: Houghton Mufflin (Revised edition).
- Windolf, P. (1979) Zur Methodologie des Interviews. In: Technische Universität Berlin (Hg.), *Soziologische Analysen*, 19. Deutscher Soziologentag in Berlin, 1979. 58

Veröffentlichungsnachweis

- 1) Die zwei Ansätze zur Erfassung der Identität und ihre Integration – Überlegungen zum Dilemma der Hochschulsozialisationsforschung. In überarbeiteter Form erschienen in: I. Sommerkorn (Hg.); Identität und Hochschule – Probleme und Perspektiven studentischer Sozialisation. Hamburg: Arbeitsgemeinschaft für Hochschuldidaktik (1981), S. 134- 152.
- 2) Kognitiv-strukturelle Analysen von Einstellungssystemen. In: H.D. Klingemann & M. Kaase (Hg.) Politische Psychologie, Sonderheft 12/1981 der Politischen Vierteljahresschrift (PVS), S. 74-86.
- 3) Reliability of Tests or Structure of Personality? Empirical Findings on the Problem of Interpreting Judgment Consistency. In: FORM-Panel Document 88. Vienna: European Coordination Center for Research and Documentation in Social Sciences.
- 4) Experimental Questionnaires: A New Approach to Personality Research. In überarbeiteter Form erschienen in: A. Kossakowski (ed.), Progress in Psychology of Personality. Berlin-Ost: Deutscher Verlag der Wissenschaften (1982), S. 132-144. (auch: Amsterdam: North-Holland).