

Georg Lind

**Der ‘Moralisches-Urteil-Test’ (m-u-t) in der Kritik
– Eine Replik auf Schmieds Analyse**

1982

Contact:

Prof. Georg Lind

University of Konstanz

FB Psychologie

78457 Konstanz

E-Mail: Georg.Lind@uni-konstanz.de

For further information and publications on this topic see

www.uni-konstanz.de/ag-moral/b-publik.htm

**Psychology of Morality &
Democracy and Education**

Der 'Moralisches-Urteil-Test' (m-u-t) in der Kritik – Eine Replik auf Schmieds Analyse

Nach einer langen Stagnation der Forschung in zentralen Bereichen der Psychologie scheint es nur folgerichtig, aufzumerken, wenn ein neuer Ansatz, wie die kognitive Entwicklungstheorie, mit dem Anspruch auftritt, einige der theoretischen und methodischen Schwierigkeiten zu überwinden, und dazu auch bereits Belege aus einer umfangreichen empirischen Forschung vorweisen kann (vgl. Broughton 1978; Kohlberg 1979; Oser 1981). Dass die zentrale Hypothese dieser Theorie vergleichsweise informativ (Popper) sind, d.h. leichter als viele andere psychologische Theorien empirisch falsifiziert werden können (Lind 1980 a), läßt es nicht nur absehbar, sondern wünschenswert erscheinen, wenn auch die deutsche psychologische Forschung diesen Ansatz rezipiert und ihn in kritischer (und kompetenter) Auseinandersetzung weiterentwickelt. Demgegenüber hat kürzlich Schmied (1981) den Wert dieser Rezeption bezweifelt. Er charakterisiert die deutsche Rezeption der amerikanischen Moralpsychologie als Aufspringen auf den 'Kohlberg-bandwaggon'. Er behauptet, dass die Rezeption "nur eine Frage der Zeit" und die Kohlberg-Theorie "doch zu verlockend" war, um nicht "forschungsmäßig genutzt zu werden" (S. 51). Schmied bemängelt, dass veraltete und durch die neuere Forschung überholte Meßverfahren wider bessere Einsicht ins Deutsche übertragen wurden (als Beispiele nennt er das 'Kohlberg-Interview zum Moralischen Urteil', KIMU, von Eckensberger et al ..., und den 'Fragebogen zum Moralischen Urteil', FMU, aus derselben Gruppe), bzw. dass deutsche Erhebungsinstrumente sich "sehr stark an amerikanische Vorbilder an(lehnen)" (S.51; hier nennt er den 'm-u-t' von Lind und seine Variante, den 'm-u-f', von Krämer-Badoni und Wakenhut). Offenbar zur Bestätigung dieser These befaßt sich der Autor kritisch mit dem 'Moralisches-Urteil-Test' (m-u-t) (S.52).

Der m-u-t ist in den letzten Jahren von einer Reihe von Forschern und Forschungsgruppen hier und in anderen Ländern übernommen oder auf der Basis einer kritischen Rezeption für eigene Forschungsfragen umgearbeitet worden (vgl. Krämer-Badoni & Wakenhut 1978; Brügelmann 1979; Katibai 1979; Schmid 1981; Briechle 1981). Wir haben in einer Reihe von Untersuchungen die Frage der theoretischen Gültigkeit dieses Tests überprüft und bestätigende Befunde erhalten (Lind 1978a; 1980a-d; Lind & Wakenhut 1980; Lind, Sandberger & Bargel 1981).

Schmieds "empirische und theoretische Analysen" des m-u-t sind die bislang gründlichsten, die von seiten eines nicht an der Konstruktion beteiligten Forschers durchgeführt wurden. Das Ergebnis dieser Analysen ist bemerkenswert, und zwar in zweifacher Hinsicht: Einmal, weil es die bisherigen Annahmen und Befunde der m-u-t Forschung erneut und ohne Einschränkung bestätigt.

Zum anderen ist es bemerkenswert, weil diese Bestätigung in der "Zusammenfassung" (S.64) von Schmied auf den Kopf gestellt und der Test dort negativ bewertet wird. Der Autor vertritt einen durch Empirie und Logik nicht zu erschütternden psychologischen Standpunkt, der weder der Kohlberg-Theorie noch dem m-u-t gerecht wird. Da solches Missverstehen auch teilweise bei anderen Autoren zu finden ist, möchte ich im folgenden nicht nur die Diskrepanz zwischen den Befunden und den Schlußfolgerungen Schmieds, sondern auch die psychologischen Grundlagen des m-u-t zu verdeutlichen versuchen.

2. Annahmen, Ergebnisse und Schlußfolgerungen der Untersuchung von Schmied – eine kritische Gegenüberstellung

Der Autor will prüfen, ob die (mit dem m-u-t verbundenen; G.L.) Ansprüche durch die vorliegenden standardisierten Meßinstrumente tatsächlich schon eingelöst sind bzw. ob sie anhand von Präferenzfragebogen überhaupt eingelöst werden können" (S. 52). Wir wollen im folgenden Punkt für Punkt seine Annahmen, Ergebnisse und Schlußfolgerungen gegenüberstellen:

- (1) Für Schmied basieren "die deutschsprachigen geschlossenen und standardisierten Verfahren ... weitgehend auf den Arbeiten von Rest [. . .] und Keasey [. . .]" (S. 51). Es ist richtig, dass der m-u-t und der m-u-f wesentliche Anregungen aus diesen Arbeiten bezogen haben. Man muß aber auch wissen, dass sich hinter dem Hinweis auf dies Autoren drei verschiedene Verfahren verbergen und dass der m-u-t selbst bei oberflächlicher Betrachtung nicht als Adaption eines dieser drei Verfahren gelten kann. Ein neues Verfahren wurde notwendig, weil diese den strukturellen Aspekt zu sehr zugunsten der üblichen psychometrischen Kriterien vernachlässigen (Lind & Wakenhut 1980, S. 326).
- (2) "Der Erwerb bzw. das Erreichen einer bestimmten Stufe der moralischen Entwicklung verläuft dabei nach Rest ... im Sinne einer zunehmenden kognitiven Verankerung von der Präferenz über das Verständnis bis zur Produktion. Die Präferenz von Argumenten einer Entwicklungsstufe würde somit das erste Signal für den Erwerb dieser neuen Stufe indizieren" (S.52). Für Schmied "scheinen vor diesem Hintergrund standardisierte Präferenzfragebogen theoretisch legitimiert" (S. 52). Mit dem (nicht von Rest, sondern von uns geprägten) Begriff der kognitiven Verankerung verbindet sich zudem die über Kohlberg hinausgehende Vorstellung, dass zwischen Präferenz und Produktion eine funktionale, die Entwicklung steuernde Beziehung besteht (Lind 1980 b).
- (3) Schmied erkennt an, dass "ein idiographischer Ansatz wie die varianzanalytische Zerlegung eines individuellen Antwortmusters gegenüber einer bloßen Betrachtung der mittleren Präferenzwerte hinsichtlich der Struktur

des moralischen Urteils einen höheren Erkenntnisgewinn (bietet)” (S.61). Der dies wieder einschränkende Hinweis auf Olweus (1976) erscheint uns unangebracht, da sich dieser (u.E. zu Recht) gegen den ‘modernen Interaktionismus’ wendet, der die individuelle Antwortvarianz auf Personen, Situationen und statistische ‘Interaktionen’ aufzuteilen versucht. Der m-u-t basiert dagegen auf dem ‘alten’ Interaktionismus von G.H. Mead (1936), für den diese Zerteilung keine akzeptable Vorstellung gewesen wäre.

- (4) Der m-u-t erhebe den Anspruch, “‘Struktur und Inhalt moralischer Urteile [...] (zu trennen)’ [...], d.h. den Entwicklungsstand der moralischen Urteilskompetenz in seinen ‘motivationalen’ (als präferierte Wertperspektiven = Präferenzurteile) und ‘kognitiven’ (als formale Struktur des moralischen Urteils) Komponenten zu erfassen” (S. 52). Dies ist in der Tat der zentrale Anspruch des m-u-t. Durch die Möglichkeit einer simultanen Messung beider Komponenten des MU ist es u.W. erstmals gelungen, die Beziehung beider Komponenten zueinander auf direktem Wege empirisch zu analysieren. Bislang waren diese entweder in einem kombinierten Stufenwert aufgegangen oder wurden durch verschiedene Testinhalte operationalisiert (vgl. Kohlberg 1969, S. 389-397; Kuhn et al 1977). Mit den Ergebnissen der Konstanzer Untersuchung konnte daher eindeutiger als bisher belegt werden, dass beide Komponenten gemäß Kohlbergs ‘Parallelitäts’-Annahme positiv miteinander korrelieren (Lind 1978 b, S. 190; 1980 d; und auch Schmied, S. 54-59).
- (5) M-u-t und m-u-ferlaubten ‘bei geeigneter Testdarbietung und -auswertung die Skalierung von Verhaltens- und Urteilskonflikten nach ihrem moralischen Aufforderungscharakter’ (Lind 1978 b, S.370), d.h., sie würden es gestatten, bereichsspezifische moralische Urteile zu erfassen” (S.52). Das ist richtig. Bestimmte Dilemmas werden von den Befragten selbst dann nicht auf der höchsten Stufe des moralischen Urteilens angegangen, wenn ihnen dieses möglich zu sein scheint. Mehr noch, entgegen Kohlbergs Theorie (vgl. 1969, S.387) wird nicht einmal immer die höchste Urteilsstufe präferiert, sondern je nach moralischem Aufforderungsgehalt der Dilemma-Situation eine darunterliegende Stufe. Diese Annahme bestätigt sich auch in den Daten Schmieds (S.55-56). Das hier angesprochene Problem der Segmentierung des MU wird derzeit intensiv untersucht (Döbert & Nunner-Winkler 1978; Lind 1978 a).
- (6) Schmied untersucht “‘im Einzelnen”, ob die Trennung der beiden Komponenten “zu einer differenzierteren Erfassung des Entwicklungsstandes der moralischen Urteilskompetenz führt als dies mit bisherigen Meßmethoden gelang” und “ob ein Präferenzfragebogen wie der M-U-T überhaupt eine hinreichende Differenzierungskraft bezüglich des Entwicklungsstandes der moralischen Urteilskompetenz besitzt. Nach allen bislang vorliegenden Ergebnissen ... erwarten wir allerdings ... selbst für eine sehr heterogene Entwicklungsgruppe (Altersspanne von 16-28 Jahren) eine nur geringe Differenzierung” (S.53). Abgesehen davon, dass diese Gruppe entwicklungsmäßig nicht allzu ‘heterogen’ ist (Schmied selbst schwankt, ob seine Vpn nun ‘homogen’, S.61, oder ‘heterogen’, S. 62, sind), ist bei dieser Frage genau zu untersuchen, was mit ‘differenziert’ gemeint sein kann: (a) Vom m-u-t wird (auch von Schmied unbestritten) behauptet, dass er ein differenziertes Bild der individuellen Struktur des moralischen Urteilerverhaltens liefert

(Lind 1978 a, S. 196). (b) Der m-u-t soll Entwicklungsunterschiede sichtbar machen. Aus der Sicht der Theorie muß hierbei klar zwischen den beiden Komponenten des MU unterschieden werden. Für die modale Präferenzstufe des MU der besagten Altersgruppen nehmen wir an, dass der Test kaum differenziert, dass vielmehr die 'bloße' Präferenz für qualitativ hochstehende Argumente bereits früh ausgebildet ist (Piaget 1973). (c) Der eigentlich bedeutsame (und für Persönlichkeitstests nicht selbstverständliche) Anspruch bezieht sich auf die Entwicklungs-Sensitivität der strukturellen Meßwerte des m-u-t. Allein darauf kann sich nur der Anspruch des m-u-t begründen, den Strukturaspekt valide operationalisiert zu haben. Schmied behauptet, dass es gerade dies mit dem m-u-t und ähnlichen Verfahren nicht möglich sei (S. 62). An anderer Stelle meint er jedoch auch: "Inwieweit die Verwendung struktureller Aspekte des moralischen Urteils wie beispielsweise Höhe des Differenzierungsgrades zwischen den einzelnen, die jeweiligen Moralstufe repräsentierenden Argumenten (Faktor 'Moralstufe des Arguments') den Einsatz des Tests rechtfertigen, bedarf noch des Nachweises" (S.61).

3. Kann mit dem m-u-t die kognitive Entwicklung des moralischen Urteilens erfaßt werden?

Wir haben drei Versuche unternommen, um diese Frage zu beantworten und ein kognitives Maß für den Stand der Moralentwicklung zu gewinnen. In einem ersten Entwurf haben wir verschiedene Strategien der Kennwertermittlung, u. a. das von Krämer-Badoni und Wakenhut (1978) diskutierte Konsistenzkriterium erprobt (Lind 1976, S. 121f). Diese Verfahren sind zwar, wie auch Schmied anerkennt, ein wesentlicher Fortschritt gegenüber der bisher üblichen Einstellungsmessung. Sie sind aber nicht völlig befriedigend, weil sie ähnlich wie das Kohlberg-Interview Inhalt und Struktur vermengen.

In einem zweiten Schritt haben wir, angelehnt an die strukturalistische Entwicklungstheorie Piagets und an die Befunde von Keasey (1974), den Grad der Orientierung des Bewertungsverhaltens an der moralischen Qualität der Argumente als Maß für die kognitive Komponente definiert (Lind 1978 a, S.189). Ermittelt wird dieses Maß durch die VKZ als Anteil des Faktors 'moralische Qualität der Argumente' (Faktor Stufe) an der gesamten individuellen Antwortvarianz (Quadratsummenanteile). Die Konstruktion dieses Maßes beruht auf der Vorstellung, dass die relevante Struktur des Bewußtseins sich in der Struktur des Urteilsverhaltens zeigt, das durch dieses Bewußtsein gesteuert wird (vgl. Kohlbergs, 1971, S.228, 'Korrespondenz-Postulat'). In der Tat zeigt das Maß der kognitiven Strukturiertheit des MU deutliche Entwicklungstendenzen. Es nimmt mit dem Alter (dem für Kohlberg zentralen Kriterium für die kognitiv-moralische Entwicklung) zu und es korreliert positiv

mit der Präferenz für höhere Stufen (Lind 1980 c; Briechle 1981; Heidbrink 1981; Schmied, S. 58-59).

In einem dritten Schritt haben wir versucht, mit rein strukturellen Maßen eine Zuordnung zu einer bestimmten Stufe vorzunehmen. Das ist bislang noch nicht gelungen, weshalb wir, viertens, als Annäherung hierzu entsprechende Bestimmungsversuche mit Aggregatdaten vorgenommen haben, und zwar mittels drei verschiedener Methoden: einer klassischen Itemanalyse, einer Kommunalitätenanalyse der Stufeninterkorrelationen und einer Stabilitätsuntersuchung des Urteilsverhaltens über zehn Monate hinweg. In allen drei Auswertungsstrategien hat sich derselbe Befund ergeben (Lind 1980 b): Oberstufenschüler und Studienanfänger scheinen kognitiv erst auf Stufe III voll verankert zu sein. Stufe IV scheint kognitiv erst von Studienabsolventen bzw. von fortgeschrittenen Studenten erreicht zu werden. Dieser Entwicklungsunterschied deutet sich auch in Schmieds Untersuchung an, insbesondere in "Tabelle 6" (S.59), in der sich ein für beide Entwicklungsstufen typischer Umschwung der Korrelationen zwischen Stufe III und IV bzw. zwischen Stufe IV und V zeigt. Auf diesen Befund konvergieren auch die neueren Auswertungen Kohlbergs (1979), nachdem die älteren Interview-ratings zu einer höheren Einschätzung des Entwicklungsstandes dieser Altersgruppen geführt hatten (Stufe IV bzw. V).

In diesen Ergebnissen sehen wir einen klaren Beleg dafür, dass die kognitiven Maße des m-u-t Entwicklungsunterschiede der moralischen Urteilskompetenz anzeigen. Offenbar bereiten aber die gegenüber herkömmlichen Einstellungs- und Persönlichkeitsfragebogen andersartige Persönlichkeitstheorie, Erkenntnistheorie und Methodologie bei der Beurteilung des m-u-t einige Schwierigkeiten.

4. Theorie und Methodologie des 'Moralisches-Urteil-Tests'

Der m-u-t ist kein Fragebogen im klassischen (d.h. in dem durch die American Psychological Association definierten und von den meisten Fachzeitschriften durchgesetzten) Sinne. Der m-u-t ist ein objektiver Verhaltens-Test zur Erfassung von Inhalt und Struktur der (sinnhaften) Interaktion von Individuum und sozialer Umwelt:

Erkenntnistheorie. Der m-u-t beansprucht allein theoretisch valide zu sein, weil er im Sinne Loevingers (1957) als "Instrument der psychologischen Forschung" konzipiert wurde. Für diesen Zweck, für die empirische Validierung von psychologischen Theorien, ist es weder notwendig noch sinnvoll, Tests empirisch zu validieren (vgl. Travers 1951). Er muß die Funktion eines theoretisch validen 'Kriteriums' ohne empirische Itemselektion

übernehmen können (vgl. Wiggins 1973, S.415; Lind 1976, S.137). Zur Sicherung seiner theoretischen Validität wurde der m-u-t in kritischer Auseinandersetzung mit der kognitiven Entwicklungstheorie und den strukturellen Persönlichkeitstheorie konzipiert und durch Experten auf die 'kategoriale Repräsentativität' der Argumente für die sechs Stufen der Moralentwicklung (nach Kohlberg 1969) hin eingeschätzt. Die zweite, derzeit gültige Version ist veranlaßt worden durch die Diskussion mit Befragten und Kollegen, die die Korrektur einiger Items und die Umstellung der sog. Bilanzfrage angeregt haben /1/. Empirisch wurde der Test nur auf die 'Interessantheit des Dilemmas' und auf die Verstehbarkeit der Items und der Instruktion hin überprüft.

Persönlichkeitstheorie. Dem m-u-t liegt das Konzept eines dynamisch-strukturellen Persönlichkeitszuges zugrunde (Lind 1980 c), wie es von G. Allport (1961) und anderen definiert wurde: *„Persönlichkeit ist die dynamische Organisation des individuellen Denkens ('mind'), welches die charakteristische Interaktion einer Person mit seiner Umwelt auf einer bestimmten Stufe der Entwicklung kontrolliert.“* Diese Definition impliziert, dass Inhalt, Struktur und Verhaltensakte zusammen eine Einheit (den 'dynamisch-strukturellen Persönlichkeitszug') bilden. Jedes ihrer Elemente ist nur durch Bezug auf die anderen definierbar. Daher können die motivationale und die kognitive Komponente nur analytisch unterschieden, aber nicht dinglich voneinander getrennt werden (Nicht-Separierbarkeits-Postulat). Schließlich ist Persönlichkeit auch nicht unabhängig vom Verhalten, sondern nur durch die Art der Interaktion mit der sozialen Umwelt erkennbar.

Methodologie. Das häufig beklagte Problem der Koordination von Theorie und Methode der Persönlichkeitsforschung haben wir durch das Konzept des 'Experimentellen Fragebogen' zu lösen versucht (Lind 1980 c). In diesem Ansatz wird die dynamische Komponente nicht als 'latentes Konstrukt' post hoc ermittelt, sondern als 'experimenteller' Faktor in die Testsituation hineinkonstruiert. Das 'moralische Urteil' ist nicht durch isolierte Verhaltensakte, sondern als 'manifestes Konstrukt', d.h. als eine empirisch verifizierbare Struktur des Urteilsverhaltens genau definiert, das ohne "verborgene anthropologische Annahmen" (Kempf 1978; Hilke 1980) auszukommen versucht. Da dieser Ansatz anhand der individuellen Antwortkonsistenzen zunächst feststellt, ob eine vermutete dynamische Disposition bei einer bestimmten Person in einer bestimmten (Test-) Situation überhaupt verhaltensbestimmend ist, gilt diese nicht, wie in der klassischen Psychometrie, als Eigenschaft des Meßinstruments, sondern der Persönlichkeit. Eine Bestimmung der 'Reliabilität' des m-u-t stünde im Widerspruch zu der ihm zugrundeliegenden Persönlichkeitstheorie.

Der m-u-t ist am ehesten noch vergleichbar mit dem theoriegeleiteten

klinischen Interview, wie es Freud und Piaget entwickelt haben. Dem Nachhaken ('probing'), mit dem im Interview die kognitive Verankerung einer moralischen Einstellung eruiert wird, entsprechen im m-u-t die Gegenargumente und, insoweit vergleichbare moralische 'issues' vorliegen, auch die Variation der Dilemma-Situation. Mit den in den Test hineinkonstruierten "situativen Schwellen" (Lind 1978 b) wird sichtbar gemacht, wieweit sich der einzelne Befragte bei der Bewertung von Argumenten sich an deren 'moralischen Qualität', am Aufforderungsgehalt der Situation und an der Übereinstimmung der Argumente mit der einmal gefaßten Meinung zum Dilemmas (Faktor 'Pro/Con') orientiert.

5. Schlußfolgerungen

Geht man die von Schmied überprüften Ansprüche des m-u-t durch, so kommt man also zu einem positiven Ergebnis. Schmieds "theoretische und empirische Analysen" bestätigen durchweg die Annahmen und Ergebnisse der m-u-t Forschung, wie er selbst mehrmals beteuert: (a) Es "zeigt sich ein für alle Altersgruppen charakteristisches Antwortmuster" (S.55). (b) Die "Daten entsprechen, was die Abfolge der Mittelwerte anbelangt, den Befunden von Rest (1973) und Lind (1978)" (S.56). (c) Das Auseinanderklaffen der Bewertung von meinungskonformen und meinungskonträren Argumenten, ein entscheidendes Kriterium für die kognitive Moralentwicklung, erwies sich in Schmieds Untersuchung als "Indiz für ein noch labiles postkonventionelles Urteilsverhalten dieser Untersuchungsgruppe" (S. 56). (d) "Lediglich auf den ersten drei Stufen des moralischen Bewußtseins deutet sich ein konsistentes Urteilsverhalten an, ein Ergebnis, das mit den Befunden von Lind (1979) übereinstimmt" (S. 56). (e) Der Entwicklungscharakter der m-u-t Werte zeigt sich auch in der Korrelation seines Strukturmaßes mit dem Alter ($r = .20$, S.57) und in der durchschnittlichen Determination der individuellen Urteilsvarianz durch das moralische Bewußtsein (Faktor 'Stufe'): es ist, wie zu erwarten war, bei Schülern deutlich niedriger (34%) als bei Studenten (46%) gewesen (die bei Lind, 1978, berichteten, insgesamt etwas geringeren Werte gehen auf die Verwendung der 1. Version des m-u-t zurück). (f) "Eine Zusammenhangsanalyse zwischen diesen Faktoren ('Stufe' und 'Pro/Con', G.L.) und den Präferenzurteilen pro Stufe liefert eine Bestätigung bisheriger Befunde" (S.58). (g) Erwartungsgemäß zeigt sich auch ein "Zusammenhang zwischen dem Faktor 'Stufe' und dem Stufen-Score aus dem KIMU" von $r = .22$ (S.58).

Dies alles widerspricht der "Zusammenfassung" Schmieds. Woanders meint er denn auch: "Die bisher mit einem Präferenzfragebogen wie dem 'Moralisches-Urteil-Test' in der Literatur präsentierten Ergebnisse erfahren in unserer

[. . .] Stichprobe in wesentlichen Punkten eine Bestätigung” (S.59). Eine solche Schlußfolgerung erscheint auch uns logisch und folgerichtig. Damit erscheint aber auch die abschätzig beurteilte Kohlberg-Rezeption in einem neuen Licht. Es dürfte sichtbar geworden sein, dass diese als Mode nur unzureichend charakterisiert ist. Wir sind auf den theoretischen und methodologischen Hintergrund des m-u-t eingegangen, um zu zeigen, dass es sich hierbei auch um eine konstruktive Kritik an dem Ansatz Kohlbergs handelt, mit der die Moralpsychologie neue Wege der Persönlichkeitsforschung geht. Wir hoffen damit auch gezeigt zu haben, dass die kognitive Entwicklungstheorie in Verbindung mit einer neuen Methodologie die Möglichkeit in sich birgt, Theorie und Methode in der Persönlichkeitspsychologie wieder enger aufeinander zu beziehen.

Zusammenfassung

Die Untersuchung von Schmied (1981) zum ‘Moralisches-Urteil-Test’ (m-u-t) wird re-analysiert. Es wird gezeigt, dass seine Befunde die theoretische Validität des m-u-t in allen überprüften Punkten bestätigen und dass deshalb die Bewertung dieses Tests in seiner “Zusammenfassung” unbegründet ist. Insbesondere bestätigt sich erneut die Entwicklungsmäßigkeit der Strukturmaße (Konsistenz des individuellen Urteilsverhaltens bezüglich einer bestimmten moralischen Orientierung). Darüberhinaus werden einige grundlegende Konzepte (theoretische Validität, dynamisch-struktureller Persönlichkeitszug und Experimenteller Fragebogen) erläutert, durch die sich der m-u-t von herkömmlichen Persönlichkeitstests unterscheidet.

Summary

Re-analyses Schmied’s (1981) investigation of the ‘Moralisches-Urteil- Test’ (m-u-t). It is shown that his findings support the theoretical validity of the m-u-t in all examined respects, and that, therefore, the evaluation of the test in his ‘summary’ is unwarranted. In particular, the developmental character of the structural measures (the consistency of an individual’s judgment behavior with regard to a particular moral orientation) is reinforced. In addition, some basic concepts (Theoretical Validity, Dynamic-Structural Personality Trait, and Experimental Questionnaires) are explained which distinguish the m-u-t from traditional personality tests.

Literatur

- Allport, G., *Pattern and growth in personality*. New York: Holt, 1961.
- Briechele, R., Die Messung der moralischen Urteilsfähigkeit in MUP und MUT. Zwei geschlossene Instrumente für Jugendliche. Arbeitsbericht 1/81, Projekt Entwicklung im Jugendalter. Universität Konstanz, 1981.
- Broughton, J.M., Criticism of the developmental approach to morality. A.P.A. Journal Supplement Abstract Service, November 1978.
- Brügelmann, H., Erfahrungen mit dem Test zum 'Moralischen Urteil' bei Jura-Studenten. In G. Lind (Hg.), *Moralische Entwicklung und soziale Umwelt*. Tagungsbericht, Universität Konstanz, 1979, 74-81.
- Döbert, R. & Nunner-Winkler, G., Performanzbestimmende Aspekte des moralischen Bewußtseins. In G. Portele (Hg.) *Sozialisation und Moral*. Weinheim: Beltz, 1978, 101-121.
- Heidbrink, H., Zur Bedeutung kognitiver Komplexität und moralischer Urteilsfähigkeit für politische Lernprozesse. FEoLL Forschungsbericht, Paderborn.
- Hilke, R., Zu den impliziten 'anthropologischen' Voraussetzungen der klassischen Testtheorie. *Diagnostica* 26, 1980, 99-111.
- Katibai, J., Präferenz für die sechs Typen moralischer Argumentation nach Kohlberg in Abhängigkeit von Geschlecht und sozialem Kontext (Bildung). Unveröff. Diplomarbeit, Universität Konstanz, 1978.
- Keasey, Ch., The influence of opinion-agreement and qualitative supportive reasoning in the evaluation of moral judgments. *Journal of Personality and Social Psychology* 30, 1974, 477-482.
- Kempf, W., Kritische Bemerkungen zu den impliziten Voraussetzungen psychologischer Testtheorie und -praxis und ihre Angemessenheit zur Erfüllung der Aufgaben curricularer Evaluation. Unveröff. Manuskript.
- Kohlberg, L., Stage and sequence: the cognitive-developmental approach to socialization. In D. Goslin (Ed.) *Handbook of socialization theory and research*. Chicago: Rand McNally, 1969, 347-480.
- Kohlberg, L., From is to ought. Th. Mischel (Ed.) *Cognitive development and epistemology*. New York: Academic Press, 1971, 151-235.
- Kohlberg, L., The meaning and measurement of moral development. Unveröff. Manuskript, Harvard University, 1979.
- Krämer-Badoni, Th. & Wakenhut, R., Möglichkeiten der Skalierung des moralischen Urteilsverhaltens. In L. Eckensberger (Hg.) *Entwicklung des moralischen Urteilens*. Saarbrücken: Universität, 1978, 379-391.
- Kuhn, D., Langer, J, Kohlberg, L. & Haan, N., The development of formal operations in logical and moral developments. *Genetic Psychological Monographs* 95, 1977, 97-188.
- Lind, G., Überlegungen und Hinweise zur Auswertung des 'Moralisches-Urteil-Tests'. In: *Arbeitsunterlage 40*, Forschungsgruppe Hochschulsozialisation. Universität Konstanz, 1976, 119-146.
- Lind, G., Wie mißt man moralisches Urteil? Probleme und alternative Möglichkeiten der Messung eines komplexen Konstrukts. In G. Portele (Hg.) *Sozialisation und Moral*. Weinheim: Beltz, 1978 a, 171-201.
- Lind, G., Der 'Moralisches-Urteil-Test': Anleitung zur Anwendung und Weiterentwicklung des Tests. In L. Eckensberger (Hg.) *Entwicklung des*

- moralischen Urteilens. Saarbrücken: Universität, 1978 b, 337-358.
- Lind, G., Zum Stand der moral judgment Forschung: Überlegungen zum Gehalt und zur Bewährung von Kohlbergs Theorie. In R. Silbereisen (Hg.) Newsletter Soziale Kognition 3, 1980 a.
- Lind, G., Zur Bestimmung des Entwicklungsstandes der moralischen Urteils-kompetenz beim Übergang vom Gymnasium auf die Universität. In J. Domnick (Hg.) Aspekte grundlagenorientierter Bildungsforschung. Forschungsbericht 18, SFB 23. Konstanz, 1980 a, 151-166.
- Lind, G., Experimental Questionnaires: A New Approach to Personality Research. In A. Kossakowski (ed.), Progress in Personality Psychology. Berlin-Ost: Deutscher Verlag der Wissenschaften, 1980 c (im Druck).
- Lind, G., Moralisches-Urteil-Test (m-u-t). Hinweise zum Einsatz und zur Interpretation eines Forschungsinstruments. Arbeitsunterlage 69, Forschungsgruppe Hochschulsozialisation. Universität Konstanz, 1980 d.
- Lind, G., Sandberger, J.-U. & Bargel, T., Moral judgment, ego-strength and democratic orientations. Political Psychology, December 1981 (im Druck).
- Lind, G. & Wakenhut, R., Erfassung von moralischem Urteil mit standardisierten Fragebogen. Diagnostica, Band XXVI, 1980, 312-334.
- Loevinger, J., Objective tests as instruments of psychological theory. Psychological Reports 9, 635-694.
- Mead, G., The philosophy of the act. Chicago: Univ. of Chicago Press, 1936.
- Olweus, D., Der moderne Interaktionismus von Person und Situation und seine varianzanalytische Sackgasse. Zeitschrift für Entwicklungspsychologie und pädagogische Psychologie 8, 1976, 171-185.
- Oser, F., Moralisches Urteil in Gruppen, soziales Handeln, Verteilungsgerechtigkeit. Frankfurt: Suhrkamp, 1981.
- Rest, J., The hierarchical pattern of moral judgment: a study of patterns of comprehension and preference of moral stages. Journal of Personality 41, 1973, 86-109.
- Schmied, D., Standardisierte Fragebogen zu Erfassung des Entwicklungsstandes der moralischen Urteilskompetenz? Diagnostica XXVII, 1981, 55-65.
- Schmid, M., Veränderungen des moralischen Urteils bei Adoleszenten. Unveröff. Lizentiatsarbeit. Universität Fribourg, Schweiz.
- Travers, R., Rational hypotheses in the construction of tests. Educational and Psychological Measurement 11, 1951, 128-137.
- Wiggins, J., Personality and prediction. Principles of personality assessment. Reading, Mass.: Addison-Wesley, 1973.

Fußnote

- 1 Der m-u-t wurde im Rahmen des Forschungsprojekts Hochschulsozialisations im SFB 23 (Universität Konstanz) entwickelt und wird in dessen international vergleichender Längsschnittstudie eingesetzt. An dem Projekt sind neben dem Autor beteiligt: Tino Bargel, B. Dippelhofer-Stiem, G. Framhein, H. Peisert (Projektleiter), J.-U. Sandberger und H.G. Walter. Für die wertvolle Hilfe bei der

Validierung des m-u-t schulde ich diesen Kollegen und auch den folgenden Dank: Wolfgang Edelstein, Thomas Krämer-Badoni, Ariane Nielsen, Gertrud Nunner-Winkler, Gerhard Portele, Ursula Schmidt und Roland Wakenhut sowie den Schülern und Studenten, die durch Ratings und als Probebefragte bei der Konstruktion des Tests mitgeholfen haben. Die Anregung für die Voranstellung der Bilanzfrage (Beurteilung der Lösung in der Geschichte) verdanken wir Lutz Eckensberger.