

MORAL IST LEHRBAR!

GEORG LIND

Mit Diskussionstheater

4. Auflage

2019, Berlin: Logos

λογος

MORAL IST LEHRBAR!

Georg Lind

4. Bedeutung und Messung der Moralkompetenz

Eine Kompetenz definiert sich durch die Art der Aufgabe, für deren Bewältigung sie benötigt wird. Moralische Kompetenz ist demnach, grob gesagt, die Fähigkeit, gemäß der eigenen moralischen Prinzipien zu handeln. Dies ist, wie jeder weiß, oft leichter gesagt als getan. Es ist zwar leicht, sich vorzunehmen, ehrlich und gerecht zu sein, keine Gewalt anzuwenden, mit anderen zu kooperieren usw. Aber es fällt uns oft schwer, immer und überall gemäß unserer moralischen Vorsätze zu handeln und zu entscheiden. Wir führen dies meist auf Willensschwäche oder auf die Umstände zurück. Wir übersehen dabei aber, dass die Schwierigkeit in der Moral selbst liegt. Wenn wir vor einem moralischen Dilemma sehen, genügt es nicht, uns mehr anzustrengen (motiviert zu sein) und die Umstände zu ändern (wenn das überhaupt immer möglich ist). Wir müssen dafür lernen, wie man eine Lösung findet.

4.1 Was ist Moralkompetenz?

Früher, in einfachen Gesellschaften, so scheint es, brauchten wir keine besonderen Fähigkeiten, um moralisch zu sein. Es reichte, wenn wir alle Regeln verinnerlichten, was man tun sollte und was man nicht tun durfte. Die Schwierigkeit bestand ‚nur‘ darin, sich diese Regeln zu merken, wenn sie sehr umfangreich wurden. Aber auch hier war man gefordert, wenn Regeln miteinander in Konflikt gerieten. Man durfte auch früher schon andere nicht töten. Aber wenn sie einen bedrohten, musste diese Regel außer Kraft gesetzt werden. Man löste solche Probleme dadurch, dass man die Gültigkeit solcher Regel auf die eigene Familie, den eigenen Stamm und später auf die eigene Nation beschränkte. Wer sich nicht an diese Regeln hielt, wurde ausgestoßen und verlor damit den Schutz dieser Regeln. Er habe „seine Rechte verwirkt“. Die Anwendung dieser Regeln führte nicht selten zu gewalttätigen Auseinandersetzungen und Kriegen.

In der Aufklärung bekam die Idee universell gültiger Regeln oder Prinzipien in der Philosophie und auch im alltäglichen Denken der Menschen Aufwind. Kant, einer der wichtigsten Vertreter der Aufklärung, definierte Moral nicht mehr als eine Liste von Verboten und Geboten für das Verhalten, sondern als die Befolgung universell gültiger Moralprinzipien. (Neiman, 2009; Höffe, 2012) Darunter verstand er solche Maximen des Handelns, von denen wir wollen können, dass sie universell gültig sind. Nach diesem, wie er ihn nannte, *kategorischen Imperativ* ist unser Verhalten also nur dann moralisch, wenn es mit unseren Prinzipien übereinstimmt. Diese neue Ethik der Aufklärung hatte aus psychologischer und pädagogischer Sicht auf den ersten Blick einen großen Vorteil. Wir müssen uns nur wenige Regeln merken. Die Befolgung der Regeln benötigt allenfalls Motivation und Mut, aber keine besondere Moralkompetenz.

Aber stimmt das? Kant postulierte, dass ein Konflikt zwischen moralischen Pflichten unvorstellbar sei. (Kant 1965, S. 24) So argumentierte er anhand eines konkreten Beispiels, dass das Prinzip der Wahrhaftigkeit immer gelten muss, und zwar auch dann, wenn man durch eine wahre Aussage das Leben eines Menschen gefährdet.² Ihm waren moralische Dilemmas nicht unbekannt. Aber offerierte uns kein Prinzip für ihre Lösung.

Im Alltag müssen wir uns der Tatsache stellen, dass gerade weil Prinzipien universelle Gültigkeit beanspruchen, sie uns immer wieder in eine Situation bringen, in der wir keine andere Wahl haben, als ein oder mehrere Moralprinzipien zu verletzen. Wie Kant scheinen viele Philosophen dieses Problem nicht gesehen zu haben.

Eine Ausnahme stellt für mich der Ansatz der kommunikativen Ethik dar, der von Jürgen Habermas (1983) und Karl-Otto Apel (1990) formuliert wurde. Soweit ich sehen kann, haben sie sich zwar nicht mit dem Problem von moralischen Dilemmas direkt auseinander gesetzt, aber ihr Ansatz bietet m. E. eine gute Grundlage für die Bewältigung von Dilemmasituationen. Ihre Ethik fordert, grob gesagt, dass wir moralische Probleme und Konflikte durch vernünftiges Abwägen von Argumenten, also durch Denken und Diskussion, lösen sollen, und nicht durch Gewalt – gegenüber uns selbst oder gegen andere –, auch nicht durch Selbsttäuschung und Betrug und auch nicht durch Macht, die auch eine Form der Gewalt, nämlich strukturelle Gewalt darstellt: Moralische Vernunft definiert Habermas (1976) als “die Tilgung jener Gewaltverhältnisse, die in die Kommunikationsstrukturen unauffällig eingelassen sind, und die bewusste Konfliktaustragung und konsensuelle Konfliktregulierung durch intrapsychische ebenso wie durch personelle Sperren der Kommunikation verhindern” (S. 34). Moralkompetenz ist also notwendig, um moralischen Dogmatismus oder Rigorismus – das heißt, das einseitige Befolgen eines Moralprinzips auf Kosten eines anderen – zu überwinden. Dogmatismus stellt auch eine Form von Gewalt dar und führt nicht selten zu manifester Gewalt oder gar zum Terrorismus. Dies, so scheint mir, ist eine moralische Maxime, die uneingeschränkt universalisierbar ist und bei der äußere Norm und interne Moral vollkommen übereinstimmen: entwickle deine Moralkompetenz!

Moralkompetenz ist noch ein relativ junges Gebiet der psychologischen Forschung und der pädagogischen Anwendung. Schon Sokrates hatte vom Vermögen gesprochen, das Gute zu tun, und Darwin (1966/1874) den Begriff der Moralfähigkeit verwendet, und hat Levy-Suhl (1912) festgestellt, dass es kriminellen Jugendlichen nicht an moralischen Prinzipien, sondern an “sittlicher Reife” fehlt. Aber erst Piaget (1973/1932) hat sich mit seinen Studien zum “moralischen Urteil des Kindes” dieses Themas systematisch angenommen. Er meint, dass man in der Entwicklung der Moral drei Phasen erkennen kann: Von der Imitation des Regelverhaltens von Erwachsenen und älteren Kindern, zur moralischen Heteronomie, das heißt der bewussten Übernahme dieser Regeln, hin zur moralischen Autonomie, das heißt der Ausrichtung des Verhaltens an eigenen Regeln. Piagets Studien fokussierten nach eigenem

² Über ein vermeintes Recht aus Menschenliebe zu lügen. (Kant-Werke Band 8, S. 638)

Eingeständnis mehr auf dem verbalen Denken von Kindern über moralisches Verhalten als auf wirklichem moralischen Verhalten. Sie befassten sich daher nur mit den Ergebnissen kindlicher Entscheidungen, und nicht mit den Denk- und Kommunikationsprozessen, die zu diesen Entscheidungen geführt haben.

Auch Kohlberg, dessen Werk u. a. auf Piaget aufbaute, studierte vorwiegend das Denken über moralische Dilemmas, aber kaum den Prozess der Lösung von Dilemmas selbst. Wie Kant verstand er unter Moralkompetenz nur die Fähigkeit, verbalisierbare moralische Prinzipien in Verhalten umzusetzen. Er definierte moralische *Urteilsfähigkeit* als

“das Vermögen, Entscheidungen und Urteile zu treffen, die moralisch sind, das heißt, auf inneren Prinzipien beruhen und in Übereinstimmung mit diesen Urteilen zu handeln.” (Kohlberg 1964, meine Übersetzung)

Vergleicht man diese Definition mit der damals (und auch heute noch) gängigen Definition von Moral als Übereinstimmung des Verhaltens mit externen Regeln und Normen, dann ist sie sehr bemerkenswert. Sie enthält drei weitreichende Neuerungen:

- dass Moral einen Fähigkeitsaspekt hat und daher nicht bloß auf ihre affektive Seite festgelegt werden kann,
- dass Moral auf *inneren* Prinzipien beruht,
- und dass Handeln Teil der Definition von Moralkompetenz sein muss. Das heißt, wir sprechen nur dann von einer Kompetenz, wenn sie sich im Verhalten zeigt, und nicht bloß als Vorstellung oder Absicht existiert.

Aber auch Kohlbergs Definition lässt einen Bezug auf die schwierige moralische Aufgabe der Bewältigung von Dilemmasituationen vermissen. Auch für Kohlberg scheint die Lösung von moralischen Konflikten und Problemen keine besondere Kompetenz zu erfordern, sondern nur eine Willensentscheidung, nämlich sich für die Umsetzung eines Moralprinzips verantwortlich zu fühlen. (Kohlberg 1984, S. 498-582) In den frühen Arbeiten zu seinem Messinstrument, dem *Moral Judgment Interview* (MJI) finden sich zwar Ansätze dazu, aber diese sind später der Revision zum Opfer gefallen. Vermutlich hat man ihre Bedeutung nicht erkannt.

Wir haben die Überlegungen von Habermas, Apel und Keasey in den Begriff und die Messung von Moralkompetenz aufgenommen. (Lind 1978; 1985a; Habermas 1983; Lind 2008a; 2012; Lind & Nowak, 2015) Wir definierten sie heute *als die Fähigkeit, Probleme und Konflikte auf der Grundlage von universellen Moralprinzipien durch Denken und Diskussion zu lösen, statt durch Gewalt, Betrug und Macht.*

Damit sollte klar sein, dass wir uns in der Psychologie und der Pädagogik nicht auf die Frage beschränken dürfen, ob Menschen gemäß ihren moralischen Prinzipien handeln können, sondern ob sie auch fähig dazu sind, die unvermeidlichen Konflikte zwischen diesen Prinzipien durch das Abwägen von widerstreitenden Gedanken und durch Kommunikation mit Gleich-

gesinnten wie auch mit Gegnern zu bewältigen. Wenn wir uns darauf beschränken würden, die Durchsetzung einzelner Moralprinzipien zu messen und zu lehren, fördern wir nicht die Moral, sondern – moralisch motivierte – Gewalt.

4.2 Zur Messung der Moralkompetenz

Als wir uns in den 1970er Jahren nach geeigneten Methoden zur Messung von Moralkompetenz umschaute, fanden wir nur einen einzigen Ansatz, der beanspruchte, Moralkompetenz zu messen, nämlich Kohlbergs *Moral Judgment Interview* (MJI). Kohlbergs “klinisches Interviewverfahren” schien uns dennoch nicht geeignet. Zum einen hatte das praktische Gründe: Das MJI war zu aufwändig in der Anwendung und Auswertung, um es in unserem damaligen Forschungsprojekt einzusetzen, in dem ca. 5000 Studierende in fünf europäischen Ländern über acht Jahre hinweg wiederholt befragt werden sollten. Zum anderen hatte unsere Entscheidung auch methodologische Gründe, die ich an anderen Stellen ausführlicher erläutere. (Lind 1978; 1985a; 1989b) Das Hauptproblem des MJI (und vieler andere Instrumente zur Messung von Moralentwicklung) sahen wir jedoch in der Messtheorie, die dem MJI zugrunde liegt.

Messtheorien sind keineswegs, wie viele glauben, ein neutrales statistisches Verfahren, das “bloß” hilft, “große Datenmengen zu reduzieren” und darstellbar zu machen. Messtheorien enthalten verborgene inhaltliche Annahmen über die Natur des Messgegenstands, die nicht immer mit unserem Wissen und unseren Annahmen über Moralkompetenz vereinbar sind. Fast alle Tests in den Sozialwissenschaften beruhen unausgesprochen auf zwei psychologischen *Annahmen* (Gulliksen 1950):

- Erstens auf der Annahme, dass sich eine Fähigkeit in *jeder einzelnen* Frage oder Aufgabe klar und eindeutig zeigt. Das heißt man nimmt an, dass die Antworten auf jede Frage (Item) eines Fragebogens – zum Beispiel das Urteil des Befragten über ein bestimmtes Argument – bis auf einen “Messfehler” vollständig von dem Messkonstrukt (hier: der moralischen Orientierung) verursacht ist. Neuere Messtheorien – wie die Rasch-Skalierung und die Item-Response-Theorie – verfeinern diese Annahme nur, stellen sie aber nicht in Frage.
- Zweitens beruhen diese statistischen Modelle für die Konstruktion und Auswertung von Tests auf der Annahme, dass der Messfehler rein zufällig ist und durch einen einfachen Trick ausgeglichen werden kann, den man der Messung in der Technik abgeschaut hat: Um den Messfehler zu verringern, misst man einfach mehrmals und berechnet dann den Mittelwert.

Beide Annahmen der klassischen Testtheorie sind unbewiesen:

- Die erste Annahme steht jedenfalls im Widerspruch zu unserer Definition von Moralkompetenz. Das einzelne Urteil ist nicht immer nur durch einen Faktor bestimmt, son-

dem meist durch mehrere gleichzeitig. Bereits Kohlberg (1958) bemerkte, dass jedes einzelne Urteil, jede einzelne Antwort, jede einzelne Handlung durch viele Faktoren gleichzeitig bestimmt sein kann: Ein moralisches Urteil "muss sowohl dem gerecht werden, was man selbst glaubt, als auch gleichzeitig der Situation. Das heißt, ein Urteil ist schwierig in dem Sinne, dass es allen Werten gerecht werden muss, die der Handelnde als wahr und wichtig ansieht" (S. 128; meine Übers.). "Ein Individuum muss konsistent über verschiedene Fragen und Ereignisse urteilen, damit man seine Stufenstruktur bestimmen kann." (Kohlberg 1984, S. 424) Wie also sollen wir aus einer einzelnen Antwort eines Befragten auf seine Moralkompetenz schließen können? Eine Kompetenz kann sich nur zeigen, wenn man die "Situation", also den Kontext der Aufgabe betrachtet und wenn man ausschließen kann, dass die Antwort durch andere denkbare Ursachen als die Fähigkeit der Teilnehmer bestimmt wird. Diese Faktoren kann man nicht einfach als Messfehler ansehen, wie die klassische Testtheorie annimmt.

- Selbst wenn es sich bei den Abweichungen der Antworten um rein zufällige "Messfehler" im statistischen Sinne handeln würde, könne man sie nicht so einfach verringern, wie dies bei der Messung lebloser Objekte meist möglich ist: Bei Menschen führt jede Messung zu Reaktionen, die die weitere Messung beeinflussen. Man kann Menschen nicht mehrmals dieselbe Frage stellen, wie es zur Reduktion eines Messfehlers notwendig wäre. Daher werden in allen Tests für Wiederholungsmessung nicht dieselben, sondern neue Fragen benutzt. Aber es gibt kaum Fragen, die verschieden sind und dennoch dasselbe messen. Daher gilt die Regel: Je länger der Test, um so geringer seine Validität. Das heißt, je länger der Test, um so genauer misst er etwas, von dem wir nicht wissen was es eigentlich ist.
- Damit eignen sich herkömmlich Tests nicht dazu, menschliche Dispositionen zu messen, schon gar nicht, um strukturelle Eigenschaften zu messen, wie es die Moralkompetenz darstellt. Strukturen, die unserem Verhalten zugrunde liegen, zeigen sich in den individuellen Mustern unseres Verhaltens. Diese individuellen Muster werden von den Testtheorien pauschal als Messfehler angesehen. "Bislang gibt es keine Theorie, die intraindividuelle Variabilität [von Antworten] erklärt" (Patry, 1989, S. 1331; meine Übers.)

Ebenso steht die Grundannahme, dass alle menschlichen Dispositionen nur "hypothetisch" und eine statistische Konstruktion sind, im Gegensatz zu unserer Vorstellung, dass sich eine Fähigkeit direkt im Verhalten zeigen muss, wenn wir sie als real akzeptieren sollen. Auch Kohlberg (1964) definierte moralische Urteilsfähigkeit als eine Fähigkeit, die sich im Verhalten zeigen muss. Ähnlich postuliert Habermas (1983), dass Kompetenzen "immer nur an ihren greifbaren Äußerungsformen, also anhand von Performanzphänomenen dingfest gemacht werden" können (S. 198).

Ursprünglich entschied sich Kohlberg für Piagets Methode des klinischen Interviews. Tests, die auf den gängigen psychologischen Messtheorien beruhen, fand er nicht geeignet, um

Moralkompetenz zu messen. (Kohlberg 1958; 1984, Kapitel 5) Später jedoch überarbeitete er und seine Mitarbeiter das MJI so, dass es den Anforderungen dieser Messtheorien genüge. (Colby et al.1987) Jetzt vertrat auch er den zentralen Standpunkt der Klassischen Testtheorie: “Meine Kollegen und ich fordern von *jeder* Frage, dass sie klar die Stufe reflektiert, der sie zugeordnet ist.” (1984, S. 403, meine Hervorhebung und Übersetzung) Ich vermute, dass er sich dem Druck beugte, den der “mainstream” auf ihn ausübte. Die revidierte Form des MJI stellt m. E. einen Versuch dar, die ursprüngliche klinische Methode mit der klassischen Testtheorie zu vereinen. Aber dies musste misslingen, da die versteckten Annahmen der Testtheorie den modernen Vorstellungen über die Struktur der Moralkompetenz diametral widersprechen. (Lind 1989b)

Bereits vor Kohlberg waren Hartshorne und May (1928, S. 11) mit der Messung des moralischen Charakters von Jugendlichen gescheitert, weil sie diese an von außen festgelegten Standards maßen. Sie haben (in ihren experimentellen Studien über das Betrugsverhalten von Kindern) Moral als die Fähigkeit definiert, der Versuchung zu betrügen zu widerstehen. Welches Verhalten als Betrug bewertet wurde, wurde also external festgelegt. Sie fanden, dass solchermaßen definiertes Verhalten stärker von situativen Bedingungen und/oder der Qualität der sozialen Erwartungen abhing als von den Dispositionen der Teilnehmer. Hartshorne und May (1928) stellten am Ende ihrer Arbeit selbstkritisch fest, dass ihre Beobachtung etwas wesentliches ignoriert hatten, nämlich die inneren moralischen Standards der Teilnehmer: “Das Wesentliche einer Handlung sind die ihr zu Grunde liegenden Absichten und diese kann daher nur beschrieben und verstanden werden, wenn man diese menschlichen Aspekte einer Situation berücksichtigt. Weder die Tat per se konstituiert den Betrug, noch allein die Intention des Handelnden, sondern die Beziehung zwischen der Handlung und seinen Intentionen und den Intentionen der Anderen” (S. 377; meine Übersetzung, GL). Auch aus philosophischer Sicht gibt es, wie wir oben gesehen haben, wichtige Einwände gegen die Definition von Moral als bloßer Übereinstimmung mit äußeren Erwartungen und Regelkonformität. (Lind & Nowak 2015)

Wenn Moralkompetenz bedeutet, dass sich ein Individuum in seinem Handeln wirklich an moralischen Prinzipien orientiert, dann reicht es nicht, wenn wir nur seine moralischen Orientierungen (Einstellungen, Werthaltungen, etc.) erfassen. Wir müssen auch die Relation zwischen Orientierungen und Verhalten analysieren. Rest (1986) schlug vor, die Entwicklung moralischer Kompetenzen mittels eines Fragebogens zu messen, der die Präferenz von Personen für eine prinzipienorientierte Moral erfasst. Der hiermit ermittelte Messwert *P* “zeigt an, in welchem Ausmaß eine Person Prinzipien-Argumente für die wichtigsten hält.” (Rest & Narvaez 1991, S. 234) Andere Forscher haben ähnliche Messinstrumente vorgeschlagen. (Gibbs et al. 1992) Moralische Orientierungen korrelieren tatsächlich oft erstaunlich eng mit moralischer Kompetenz, so dass wir mit Piaget von einer affektiv-kognitiven Parallelität sprechen können (Lind 2002). Aber diese Parallelität lässt sich nur in bestimmten Kontexten nachweisen. Sie kann, wie die Simulationsstudien von Emler zeigen, sich völlig auflösen, da Menschen ihre moralischen Orientierungen bei Bedarf in jede Richtung simulieren können.

(Emler et al. 1983; Emler et al. 1998) Daher kann man auch leicht mit Hilfe von Vorträgen über Ethik die moralischen Einstellungstestwerte von Lernenden beeinflussen. (Penn 1990) Es ist allerdings fraglich, ob die Änderung dieser Testwerte wirklich eine Veränderung der moralischen Orientierungen der Lernenden anzeigt, oder – wie im Experiment von Emler – nur ihre Fähigkeit, erwünschte Antworten zu geben (Lind 2002; siehe auch Kapitel 4.4.1 unten).

Zwei weitere Beispiele zeigen, wie Moral und Moralkompetenz nicht gemessen werden können. In den Studien von Lennick und Kiel (2005) sowie von Wang et al. (2015) werden Moral bzw. Charakter mit jeweils eigenen Selbstbeschreibungs-Fragebogen zu erfassen versucht. Schauen wir uns eine Studie näher an. Um die Tugend ‚Ehrlichkeit‘ (*trustworthiness*) zu erfassen, gaben Wang und seine Kollegen Aussagen vor wie “Man kann sich darauf verlassen, dass ich die Wahrheit sage”. Als Antwort konnten die Teilnehmer auf einer 5-Punkte Skala ankreuzen “Das bin ich überhaupt nicht” bis “Das bin genau ich”. Zwar scheint es Menschen bis zu einem gewissen Grad möglich zu sein, ihre eigene Moralkompetenz einzuschätzen. Aber ob sie auf solche Fragen objektiv antworten, ist sehr fraglich. Wang und Kollegen (2015, S. 190) fanden auch keine Bestätigung ihrer Hypothese, dass der Charakter von Jugendlichen sich durch ihre Aktivitäten bei den Pfadfindern verbessert. Bei Nicht-Pfadfindern fanden sie sogar eine Verringerung der Testwerte mit dem Alter. Dies deutet m. E. stark darauf hin, dass ihr Instrument keine Moral, sondern die Anpassung der Teilnehmer an die (vermeintlichen) Erwartungen der Forscher misst: mit dem Alter nimmt diese Anpassung gewöhnlich ab, nur nicht bei Pfadfindern, die zu Gehorsam gegenüber Autoritäten erzogen werden.

Keines der bisherigen Instrumente ist aus einem oder mehreren Gründen geeignet, Moralkompetenz zu erfassen:

- Tests, bei denen die Moralität des individuellen Verhaltens in Bezug auf externe Normen definiert und gemessen werden, statt in Bezug auf die internen moralischen Orientierungen, messen den Grad der Normkonformität einer Person, aber nicht ihre Moralkompetenz.
- Testwerte, die sich beliebig in jede Richtung simulieren lassen – wie moralische Einstellungstests –, zeigen keine Fähigkeit an. Damit sich bei Menschen Fähigkeiten überhaupt zeigen, muss der Test eine moralische Aufgabe enthalten, deren Bewältigung Moralkompetenz erfordert. Kompetenz können nach unten, aber nie nach oben simuliert werden.
- Man kann auch Moralkompetenz nicht erfassen, indem man Menschen einfach nach ihrer Moralkompetenz fragt. Zum einen haben sie davon oft nur eine vage Vorstellung davon; zum anderen versuchen die meisten, das zu antworten, was den vermeintlichen oder wirklichen Erwartungen des Forschers (als Repräsentant der Gesellschaft) entspricht. Diese Tests messen also eher soziale Erwünschtheit als Moralkompetenz.

Wie also misst man moralische Kompetenz, fragte ich mich, als ich vor vierzig Jahren den Auftrag bekam, ein geeignetes Instrument auszuwählen (Lind 1978)? Die Antwort auf diese Frage stellte der “Moralische Kompetenz-Test” (MKT) dar, der seit 1977, von redaktionellen Änderungen abgesehen, bis heute unverändert benutzt wird. Nur den Namen habe ich vor einiger Zeit geändert, weil wir eigentlich kein vergängliches Phänomen wie einzelne *Urteile* damit messen wollten, wie der anfängliche Name moralisches Urteil-Test” (MUT) suggerierte, sondern eine stabile menschliche Disposition wie Kompetenz.³

4.3 Der *Moralische Kompetenz-Test* (MKT)

Was Not tat war die Entwicklung einer neuen Messtheorie, die mit den Annahmen der modernen Moralpsychologie übereinstimmt (Kapitel 3). Wenn wir Moral als Befolgen von inneren moralischen Prinzipien definieren, dann dürfen wir die moralische Kompetenz von Menschen nicht daran ablesen, ob ihr Verhalten mit äußeren Erwartungen und sozialen Normen und Regeln übereinstimmt, wie das von allen gängigen Testtheorien verlangt wird. Wir haben daher eine von Grund auf neue Messtheorie entwickelt, bei der wir uns an den Methoden der experimentellen Psychologie orientiert haben. (Lind 1978; 1985a; 2008b)⁴

Gemäß unserer Theorie ist die grundlegende moralische Tugend nicht die Fähigkeit, sich an externen Erwartungen zu orientieren, sondern die Fähigkeit, sich an eigenen Moralprinzipien zu orientieren und moralische Konflikte und Probleme durch Denken und Diskussion zu lösen, also sich mit Argumenten auseinander zu setzen und Prinzipien gegeneinander abzuwägen. Diese Fähigkeit ist, wie schon Sokrates beobachtete, bei uns Menschen oft nur schwach ausgebildet. Keasey (1973) zeigte in seinen Experimenten, dass es Menschen schwer fällt, sich mit kontroversen Argumenten auseinander zu setzen. Das zeigen auch unsere Erfahrungen im Alltag. Wenn wir mit Anderen eine kontroverse Diskussion über einen Sachverhalt führen, werden wir oft nach kurzer Zeit “persönlich”. Wir streiten dann

³ Unter dem Einfluss von Piaget und Kohlberg sprach ich früher abwechselnd von moralischem *Urteil* und moralischer *Kompetenz*. Eine wissenschaftlich belanglose, aber wahre Geschichte ist, dass ich einmal träumte, dass Kohlberg, dem ich von meinen Bemühungen erzählte, einen geeigneten Namen für meinen neuen Test zu finden, antwortete: “Nur MUT!” Die Zweideutigkeit dieser Antwort hatte mich so stark zum Lachen gebracht, dass ich davon aufwachte.

⁴ Der MKT (englisch: MCT, Moral Competence Test) ist kein psychometrischer Test im klassischen Sinne, sondern ein objektives Verhaltens-Experiment, mit dem Hypothesen über die Wirksamkeit bestimmter Dispositionen im individuellen Urteilsverhalten getestet werden können. Der MKT hat ein 2 x 2 x 6 faktorielles multivariates Design, also insgesamt 24 Argumente (Items). Durch dieses Design des MKT kann man am Antwortmuster einer Person ablesen, wie stark sie sich beim Ausfüllen des MKT an jedem der drei Faktoren (moralische Qualität der Argumente, Meinungskonformität und Kontext) orientiert hat.

nicht mehr über die Sache, sondern über einander, werden laut oder kehren uns den Rücken zu.

Auf diesem Hintergrund wurde diese Idee geboren, Moralkompetenz von Menschen dadurch zu messen, dass wir beobachten, wie sie die Aufgabe lösen, sich mit Argumenten auseinanderzusetzen, und zwar besonders mit Argumenten, die dem eigenen Standpunkt widersprechen. Im Folgenden will ich die Grundidee des MKT darlegen, die ich an anderer Stelle schon ausführlich beschrieben habe. (Lind 1978; 1985a; 2008b; 2015b)

Aufbau des MKT

Beim MKT werden die Teilnehmer mit zwei Dilemmageschichten konfrontiert (Selbstjustiz/Arbeiter, Sterbehilfe/Arzt), in denen die Protagonisten eine Entscheidung treffen. Die Teilnehmer bewerten diese Entscheidung auf einer Skala. Im Anschluss daran präsentierten wir ihnen *Argumente für und gegen ihre Entscheidung*. Jedes einzelne Argument wurde so gewählt bzw. konstruiert, dass es jeweils eine bestimmte Ausprägung von drei möglichen Orientierungs-Faktoren repräsentierte, nämlich 1. einen bestimmten Dilemma-Kontext (Arbeiter, Doktor), 2. eine bestimmte Haltung gegenüber der Entscheidung des Protagonisten (Pro oder Contra) und 3. einen bestimmten Typ von moralischer Orientierung.

Beispiel: Das Argument “Der Arzt hat falsch gehandelt, der krebserkrankten Frau beim Sterben zu helfen, weil das vom Gesetz verboten ist”

- bezieht sich auf einen bestimmten Kontext (“Sterbehilfe”),
- steht zudem für die Ablehnung der Entscheidung des Arztes und
- repräsentiert auch noch eine bestimmte moralische Orientierung (Kohlberg-Typ 4: Orientierung an Gesetz und sozialer Ordnung).

Jeder Teilnehmer soll für jedes Argument sagen, wie er es beurteilt, indem er auf einer Skala von -4 (lehne völlig ab) bis +4 (stimme völlig zu) eine Antwort ankreuzt.

Unsere langjährigen Erfahrungen mit dem MKT in sehr vielen Studien zeigen, dass Teilnehmer ihn meist gern ausfüllen, weil, wie manche sagen, er nicht so “langweilig” sei wie andere Fragebögen. Nur in wenigen Fällen stieß der MKT auf Unverständnis. Einmal empfand ein Befragter die Beurteilung von Argumenten als überflüssig: “Ich habe doch angekreuzt, was ich von der Entscheidung des Arztes halte! Wieso muss ich meine Antwort mehrmals wiederholen?” fragte ein Teilnehmer. Offenbar war ihm die Unterscheidung zwischen Meinung und Argument fremd. Nur in wenigen Fällen werden Fragen ausgelassen. Es scheint, dass dies meist versehentlich geschieht. In der elektronischen Version kann man den Teilnehmer daran erinnern, seine Antworten auf Vollständigkeit zu prüfen, bevor er sie versendet.

Damit die Argumente eine unterschiedliche *moralische Qualität* repräsentierten, haben wir der Auswahl die Typologie moralischer Orientierungen von Kohlberg (1984) zugrunde ge-

legt. Er unterscheidet sechs Typen der moralischen Orientierung, die im Kasten auf Seite 71 stichwortartig dargestellt sind. Kohlberg hat diese Typen der Moralorientierung in leicht veränderter Form mehrfach publiziert und dabei zeitweise auf die Stufe 6 verzichtet. In den späten Publikationen taucht sie wieder auf. (Kohlberg et al. 1986)

Die Argumente für den MKT haben wir größtenteils den Interview-Beispielen im Auswertungshandbuch für das *Moral Judgment Interview* (Colby, Kohlberg et al. 1987b) entnommen. Wenn wir keine geeigneten Beispiele fanden, hatten wir selbst Argumente formuliert. Entgegen unserer Erwartung fiel es uns meist schwerer, Argumente einer niedrigen Stufe zu formulieren als Argumente einer höheren Stufe. So kostete es uns große Anstrengung, Argumente für und gegen Sterbehilfe vom Typ 1 und 2 so zu formulieren, dass sie nicht ironisch klangen. Das hat sicher damit zu tun, dass unser moralisches Bauchgefühl infolge der intensiven Beschäftigung mit diesen Themen stark sensibilisiert war. Alle Argumente wurden in Voruntersuchungen getestet. Um sicherzustellen, dass die Argumente den sechs Typen möglichst eindeutig entsprechen, hatten wir mehrere Kenner der Kohlberg-Theorie gebeten, die Argumente den Typen zuzuordnen. (Lind 1978; 1985a)

Die Moralkompetenz eines Testteilnehmers zeigt sich in dem *Muster seines Antwortverhaltens*. Von Moralkompetenz sprechen wir dann, wenn das Muster der individuellen Antworten erkennen lässt, dass der Teilnehmer fähig ist, die vorgegebenen Argumente nach ihrer moralischen Qualität zu beurteilen statt danach, ob die Argumente mit der eigenen Meinung zum geschilderten Fall übereinstimmen oder ihr widersprechen.

Messwerte für moralische Orientierungen und für Moralkompetenz

Moralische Orientierung: Die Werte für die sechs moralischen Orientierungen eines Teilnehmers werden durch einfache Mittelwerte abgebildet: Die Beurteilungen der vier Argumente, die eine moralische Orientierung repräsentieren (zwei für jede Dilemmageschichte, je eins für und eins gegen die Entscheidung des Protagonisten) werden zusammengezählt und durch vier geteilt. Dadurch ergeben sich sechs Messwerte für moralische Orientierungen. Jeder reicht (wie die Antwortskalen) von -4 bis +4. Manchmal werden die Summenskalen nicht durch vier dividiert, so dass die Messwerte von -16 bis +16 reichen können. Die sechs Orientierungswerte werden meist als Profile dargestellt (siehe Abbildung auf Seite 55).

Moralkompetenz: Der Messwert für die Moralkompetenz wird aus dem gesamten *Urteilmuster* einer Person gebildet, das aus 24 Einzelurteilen besteht. Dazu wird zunächst errechnet, wie stark die 24 Antworten eines Befragten streuen, das heißt, wie groß die individuelle Antwortvariation ist. Dann wird errechnet, wie groß der Teil der Variation ist, die auf die moralische Qualität der beurteilten Argumente zurückgeht, und zu der gesamten Streuung seiner Antworten ins Verhältnis gesetzt. Das Resultat wird mit 100 multipliziert, damit sich ein Messwert zwischen 0 und 100 ergibt. Dieser Messwert, der C-Wert, zeigt also an, wie stark sich ein Teilnehmer an der moralischen Qualität der Argumente orientiert, statt an ihrer Meinungskonformität oder anderer Faktoren. (Lind 1978; 1985a) Es hat sich eingebürgert, den C-Wert graphisch immer von 0 bis 40 darzustellen, außer wenn höhere Werte als 40 vorkommen. Durch diese einheitliche Darstellung fällt es dem Leser leichter, die Ergebnisse verschiedener Studien miteinander zu vergleichen. In fast allen Studien liegt der mittlere C-Wert im Bereich zwischen 0 und 40. Höhere C-Werte sind selten, was zeigt, dass das Ausfüllen des MKT für viele Menschen eine schwierige Aufgabe darstellt.

Der C-Wert wird in jeder Untersuchung immer gleich gemessen und ist daher prinzipiell über alle Studien hinweg vergleichbar. Das gilt auch für internationale Studien, da Übersetzungen des MKT immer einer strengen und aufwändigen Validitätsprüfung unterzogen werden müssen, bevor sie als valide zertifiziert werden. Um sicher zu stellen, dass die Messwerte bedeutungsgleich mit dem Originaltest sind, müssen strenge Anforderungen erfüllt werden. Es reicht nicht aus, dass die Argumente korrekt übersetzt sind. (Lind 1986a; 2008b) Wir kennen alle das Problem, wenn wir eine fremde Sprache sprechen. Auch wenn wir die Sprache gut

Typen der moralischen Orientierung nach Lawrence Kohlberg.

Typ 1	Das Motiv, eigene (körperliche) Schäden und Verletzungen ('Strafe') zu vermeiden.
Typ 2	Das Motiv, Vorteile und Vergünstigungen ('Belohnungen') zu erlangen, auch wenn dabei gelegentlich Nachteile in Kauf genommen werden müssen.
Typ 3	Das Motiv, Missbilligung durch andere zu vermeiden bzw. ihre Anerkennung zu erreichen.
Typ 4	Das Motiv, die Gesetze und die Ordnung der Gesellschaft zu achten und zu ihrer Aufrechterhaltung beizutragen.
Typ 5	Das Motiv, die Regeln freier, demokratischer Willensbildung (Mehrheitsprinzip etc.) zu beachten und die auf ihnen gründenden Verträge einzuhalten.
Typ 6	Das Motiv, die Würde jedes einzelnen Menschen zu achten und sein Handeln an den universellen Prinzipien der Gerechtigkeit, Vernunft und Logik zu messen.

beherrschen, haben wir doch oft Probleme mit der Vieldeutigkeit von Begriffen. Eine Übersetzung kann völlig korrekt sein, und dennoch für unsere Zuhörer eine andere Bedeutung haben als die, die wir meinen. Mit den Validitätsstudien kann man gezielt Mehrdeutigkeit in Übersetzungen aufspüren, auch wenn man die Sprache nicht beherrscht. Meist erkennen dann die Übersetzer sofort, wie ein Argument geändert werden muss, damit es bedeutungsgleich mit dem Original ist. Bevor solche Änderungen akzeptiert werden, müssen sie nochmals in einer Validierungsstudie überprüft werden.

Obwohl der MKT die individuelle Moralkompetenz misst, sollten den Werten von Einzelpersonen keine zu große Beachtung geschenkt werden, da wir nicht wissen, wie stark sie situativ schwanken. In der Forschung und bei Wirkungsstudien werden gewöhnlich Messwerte von mehreren Testpersonen (es sollten mindesten fünfzehn sein) gemittelt, so dass das Ergebnis meist hinreichend stabil und genau ist, um Unterschiede von einem C-Punkt interpretieren zu können. Das sind Erfahrungswerte nach vielen Jahren der Forschung, nicht logisch beweisbare Werte. Ich denke, wir sind hier eher zu streng als zu liberal. Vielleicht ist das auch der Grund, warum sich die meisten Ergebnisse der MKT-Forschung gut replizieren lassen.

Der MKT darf nicht für die Bewertung und Selektion von Menschen benutzt werden. Zum einen, weil er keine genaue Beurteilung von Personen zulässt. Zum anderen aber auch, weil der Test sehr schnell korrumpiert werden würde. Wie bei anderen Selektionstests auch, würde viel Energie darauf verwendet, im MKT zu betrügen. (Campbell 1976) Nach einer Weile würde der Test dann die Bemühungen der Teilnehmer messen, ihn zu korrumpieren, aber nicht mehr die moralische Kompetenz. Er könnte dann nicht mehr für seriöse Forschung verwendet werden. Immerhin wird der MKT schon vierzig Jahre lang unverändert eingesetzt, was für ein psychologisches Instrument eine lange Zeit ist.

Einsatz des MKT

Der MKT kann als Gruppentest (z.B. in einer Schulklasse) oder als Einzeltest oder online im Internet dargeboten oder postalisch verschickt werden. Forscher sollten in ihren Publikationen von MKT-Ergebnissen möglichst immer auf die Art des Einsatzes hinweisen, da sie auf den C-Wert Einfluss haben können. Nicht betroffen davon sind aber Vergleiche und Analysen innerhalb einer Studie. Auch die Ausfallrate sollte berichtet werden. Wenn viele Teilnehmer ausfallen, steigt der C-Wert der Untersuchungsgruppe meist, da die Ausfallrate mit der Moralkompetenz korreliert: Je höher die Moralkompetenz, desto eher beteiligen Leute sich an Befragungen, wenn sie dafür ausgewählt wurden. (Krebs & Rosenwald 1977)

Der MKT erfordert grundlegende Lesefähigkeit. Das Ausfüllen der Standardversion mit zwei Dilemmas dauert je nach Lesegeschwindigkeit der Teilnehmer zwischen acht und fünfzehn

Minuten. Für den MKT werden *keine* Zeitgrenzen vorgegeben, aber bei überlanger Beschäftigung mit den Antworten wird zur schnelleren Bearbeitung ermuntert.

Der MKT kann in der Regel bei Personen ab dem Alter von ca. acht Jahren eingesetzt werden. Bei Kindern und Jugendlichen und bei Personen mit geringen Erfahrungen mit dem Ausfüllen von Fragebogen und/oder geringem Bildungsstand kann die Darbietung des Tests etwas erleichtert werden:

- Die Schrift kann vergrößert und der Test über mehr Seiten verteilt werden, aber möglichst so, dass immer sechs Argumente auf derselben Seite zu stehen kommen.
- Die Antwortskala kann von neun Punkten auf fünf verringert (von “-2” bis “+2”) werden.
- Der MKT kann individuell oder in kleinen Gruppen dargeboten werden, damit der Testleiter einzelnen Teilnehmern unbekannte Worte erklären kann. Natürlich darf dabei der Teilnehmer nicht in seinen Antworten beeinflusst werden.

Wir nehmen an, dass diese Änderungen keinen Einfluss auf den Testwert haben. Vergleiche zwischen Untersuchungen mit unterschiedlichen Testformen sollten dennoch vorsichtig interpretiert werden. In Publikationen über solche Studien sollte auf Abweichungen vom Standard-MKT immer hingewiesen werden.

Fähigkeitstests wie der MKT können empfindlich auf bestimmte äußere Umstände der Testanwendung reagieren. Folgende Umstände können zu mehr oder weniger großen Verringerungen des C-Wertes führen, der beim MKT die Moralkompetenz anzeigt:

- Zeitdruck kann zu niedrigeren Werten führen. Daher wird der MKT, wie schon gesagt, ohne Zeitbeschränkung durchgeführt. Nur wenn Teilnehmer exzessiv Zeit benötigen, werden sie ermutigt, weniger lang zu überlegen.
- Ermüdung. Wenn die Teilnehmer mehrere Fragebogen ausfüllen müssen, soll der MKT immer möglichst weit am Anfang bearbeitet werden.
- Äußere Erwartungen. Beim MKT gibt es keine äußeren Kriterien für “richtig” oder “falsch”; vielmehr wird die individuelle Moralkompetenz an den moralischen Orientierungen des Einzelnen gemessen. Wenn die Umstände den Teilnehmern signalisieren, dass es sich um eine “Leistungsprüfung” handelt, sinkt der C-Wert meist. Worin dieses Signal besteht, muss vom Untersuchungsleiter jeweils herausgefunden werden. In einer Studie ging dieses Signal offenbar von einem Intelligenztest aus, den die Befragten vor dem MKT ausfüllen mussten.

Vergleiche über verschiedene Studien hinweg können daher zu falschen Schlussfolgerungen führen, vor allem dann, wenn die Autoren der Studien es versäumen, den Leser auf diese Umstände aufmerksam zu machen. Nicht betroffen davon sind Vergleiche innerhalb derselben Studie. Für die Anwendung, Auswertung und Interpretation des MKT sind generell gute moralpsychologische Kenntnisse notwendig. (Lind 2008b)

4.4 Welche Bedeutung hat Moralkompetenz für das “Verhalten”?

Ich habe den Begriff “Verhalten” in Anführungszeichen gesetzt, weil er vielen klarer erscheint als er in Wirklichkeit ist. Wir sollten daher klären, was wir darunter verstehen. Der Begriff ist so unbestimmt wie das Wort “Ding”. Im Alltag wissen wir sofort, was jemand damit meint, wenn er gleichzeitig auf das Ding zeigt. Aber auf Verhalten kann man oft nicht einfach zeigen. Verhalten wird daher auch im Alltag oft missverstanden. Wenn Inga zu Heinz sagt “Dein Verhalten war nicht in Ordnung”, dann nimmt Heinz an, dass Inga das Verhalten meint, das er eben gerade gezeigt hat. Aber Inga kann ein anderes Verhalten von Heinz gemeint haben, nämlich das von gestern. Oder: Was für Inga nicht in Ordnung war, kann für Heinz völlig OK gewesen sein. In der weiteren Konversation kann sich herausstellen, dass Inga das Verhalten von Heinz als Ausdruck einer bestimmten Absicht sah, Heinz aber bestreitet, dass er diese Absicht hatte.

Diese Vieldeutigkeit des Begriffs *Verhalten* führt nicht nur im Alltag zu Problemen, sondern auch in der Moralforschung und der Moralerziehung, wobei Wissenschaftler und Lehrer meist den Kontext des Verhaltens ihrer Teilnehmer noch wenig kennen. Wenn sie Verhalten einordnen möchten, müssen sie auf Annahmen über menschliches Verhalten zurückgreifen, die stimmen können oder auch nicht. Verhalten ist, wie wir im Zusammenhang mit der Messung von Moralkompetenz gesehen haben (Kapitel 4.2), meist das Ergebnis eines komplexen Entscheidungs- und Abwägungsprozesses, der nicht bewusst ablaufen muss und über den wir uns daher oft selbst nicht völlig im Klaren sind. Verhalten bedeutet oft etwas anderes je nachdem, ob wir selbst der Handelnde sind, oder ob wir das Verhalten eines Anderen beobachten. Die Absichten, die wir mit dem eigenen Verhalten verknüpfen, müssen nicht von Anderen so verstanden werden, wie wir sie meinen. Demnach können die Kriterien, an denen wir moralischen Verhalten messen, ganz verschieden sein, wenn wir es an “inneren Standards”, das heißt an den moralischen Orientierungen des Individuums, oder an “äußeren Standards”, das heißt an den moralischen Erwartungen anderer Menschen messen – die zudem oft nicht einheitlich sind. Ich denke, es wäre schon viel gewonnen, wenn wir bei der Frage nach der Verhaltensrelevanz von Moralkompetenz zwischen der Innen- und der Außenperspektive unterscheiden würden, das heißt, zwischen dem Verhalten, wie es sich aus der Sicht des Handelnden darstellt und dem Verhalten, wie Andere (auch wir als Beobachter und Pädagogen) es sehen.

Mancher wird sich fragen, warum wir nicht auch der Frage nachgehen, wie sich moralische Orientierungen (Einstellungen, Werthaltungen etc.) auf das Verhalten auswirken. Nun der Grund ist, wie wir oben gesehen haben, einfach. Orientierungen sind zweifellos sehr wichtig für das moralische Verhalten: ohne sie würde es kein moralisches Verhalten geben. Aber sie können keine Unterschiede im Moralverhalten erklären, weil alle Menschen im Grunde – von geringen Variationen abgesehen – die gleichen moralischen Orientierungen haben (siehe Kapitel 3). Statistiker wissen: “No variation – no correlation”.

4.4.1 Wie zeigt sich Moralkompetenz im eigenen Urteilsverhalten? (Innenperspektive)

Die Vermutung, dass es so etwas wie moralische Kompetenz geben muss, wird durch die Beobachtung nahegelegt, dass zwischen unseren moralischen Orientierungen und unserem Verhalten in konkreten Situationen oft eine große Kluft besteht. Aber gibt es Moralkompetenz wirklich? Lässt sie sich im Urteilsverhalten von Menschen nachweisen? (Eine andere Frage ist es, ob sich Moralkompetenz auch im sozialen Verhalten nachweisen lässt, also in dem Verhalten, das andere von uns erwarten. Hierauf werde ich unten näher eingehen.)

Die Frage, ob das Urteilsverhalten von Menschen überhaupt von ihren moralischen Orientierungen beeinflusst wird, lässt sich noch genauer formulieren, da Forscher zu dieser Annahme einige konkrete Hypothesen formuliert haben:

5. Mit Sokrates nehmen wir an, dass das Vermögen, das Gute zu tun (er nannte es "Tugend", wie nennen es Moralkompetenz) im Unterschied zu dem Wollen des Guten, das heißt in unserer Terminologie, zu den moralischen Orientierungen sehr breit gestreut ist. Tatsächlich zeigt sich in allen Studien mit dem *Moralische Kompetenz-Test* (MKT, siehe oben, Kapitel 4) eine sehr breite Streuung der C-Werte, des Indikators für Moralkompetenz.
6. Alle Menschen haben hohe moralische Orientierungen (siehe oben Kapitel 3). Selbst Kriminelle haben hohe moralische Ideale. (Levy-Suhl 1912; Nucci 1995, S. 25) Daran kann es also kaum liegen, dass es in der Welt so viel Gewalt, Betrug, Korruption und Machtmissbrauch gibt, weil die Orientierungen der meisten Menschen moralischer kaum sein können.
7. Wenn Moral im Kern eine Fähigkeit ist und nicht simuliert werden kann (siehe unten), dann müssen wir annehmen, dass sie sich nur durch gute Bildung vergrößern lässt. Tatsächlich korreliert von allen untersuchten Variablen die Länge und Qualität der Bildung, die Menschen bekommen haben, am stärksten mit dem Niveau ihrer Moralkompetenz. (Lind 2002; Rest & Thoma 1985; Rest & Narvaez 1991) Die meisten Korrelationen mit anderen Variablen wie Alter, sozialer Herkunft, Geschlecht und Kultur sind viel niedriger oder werden sehr klein, wenn man das Niveau der Bildung berücksichtigt, die den eigentlichen Unterschied ausmacht. (Lind 2002; 2015a; siehe auch Kapitel 5) Der Befund, dass die Messwerte mit dem Kohlberg-Interview (MJI) stärker mit dem Alter als mit dem Bildungsabschluss korrelieren, kann daran liegen, a) dass dieser Test für die Korrelation mit dem Alter optimiert wurde (Colby, Kohlberg et al. 1987), b) dass der Bildungsabschluss allein wenig über die Qualität der Bildung aussagt, und c) dass das MJI mehr die moralische Orientierung der Probanden misst als ihre Moralkompetenz. (Lind 1989b)

8. Schließlich sollte sich die Fähigkeitsnatur der Moral darin zeigen, dass sie sich nicht einfach nach oben simulieren lässt, wie das bei moralischen Einstellungen möglich ist. Fähigkeiten lassen sich nur nach unten schwindeln, aber nicht nach oben.

Auch die letzte Annahme hat sich experimentell belegen lassen. Ich will hier etwas näher darauf eingehen, da sie von zentraler Bedeutung für das Konzept der Moralkompetenz ist. Die Idee zu einem geeigneten Versuchsaufbau zur Beantwortung dieser Frage stammt von Emler und seinen Kollegen (Emler et al. 1983), die damit die zentrale Annahme der Kompetenz-Theorie der Moral widerlegen wollten. Ihr Versuchsaufbau war wie folgt: Die Versuchspersonen, Studierende, wurden zunächst gemäß ihrer eigenen Angaben in drei Gruppen aufgeteilt: mit linken, neutralen und rechten politischen Haltungen. Dann wurde ihnen ein moralischer Einstellungstest, der schon erwähnte *Defining-Issues-Test* (DIT) von Rest (1986) zum Ausfüllen vorgelegt. Der DIT enthält wie der MKT eine Reihe von Dilemmageschichten. Aber anders als im MKT werden den Befragten keine Argumente für und gegen die eigene Meinung vorgelegt, sondern Aussagen ("Issue statement"), die verschiedene Kohlbergstufen thematisieren. Diese sollten von den Befragten nach ihrer Wichtigkeit bewertet und in eine Rangreihe gebracht werden. Aus den Antworten wird dann für jede Person ein so genannter P-Wert berechnet, der von 0 bis 95 variieren kann, und der anzeigt, für wie stark sie moralische Prinzipien vor anderen Orientierungen bevorzugt. (Rest 1979; Rest et al. 1999)

Dieser erste Teil des Versuchs zeigte, dass die moralische Einstellung (P-Wert) hoch mit der politischen Grundhaltung der Versuchspersonen korreliert: "linke" Teilnehmer erhielten deutlich höhere P-Werte als "rechte." (Gielen 1986) Aufgrund dieses Befunds sahen sich Emler und Kollegen in ihrer Vermutung bestätigt, dass es sich bei der Moral um eine Einstellung handelt, ähnlich einer politischen Einstellung. Um wirklich zu zeigen, dass es sich nicht um eine Fähigkeit handelt, führten sie noch einen zweiten Versuch durch. Dieselben Versuchspersonen wurden gebeten, sich in die politisch entgegengesetzt eingestellte Personen hineinzusetzen und den DIT so auszufüllen, wie diese das tun würden. Wenn es den Versuchspersonen möglich sein würde, ihr Urteilsverhalten in jeder gewünschten Richtung zu simulieren, dann sei der Beweis erbracht, dass es sich dabei nicht um den Ausdruck einer Fähigkeit handeln kann. Besonders kritisch ist natürlich die Simulation von Testwerten, die höher liegen als die eigenen, also die Simulation der Moralkompetenz der "Linken" durch die "Rechten".

Das Ergebnis des Emler-Experiments wird in der Abbildung auf Seite 77 dargestellt. Es zeigt, dass es in der Tat möglich ist, den DIT-Messwert nach oben zu simulieren. Die Gruppe der "konservativen/rechten" Studenten, die in der ersten Runde niedrige P-Werte erzielten, können ihren P-Wert auf das Niveau der "linken" Studenten steigern, wenn man sie darum bittet. Dass die "linken" Teilnehmer ähnlich gut ihre Werte nach unten simulieren konnten, ist nicht überraschend. Emler et al. (1983) zogen aus diesem Ergebnis zwei Schlüsse: a) Es gibt keine moralischen "Fähigkeiten" und b) der DIT ist nicht geeignet, solche Fähigkeiten

zu messen. Damit schien bewiesen, dass Moral bloß eine Sache der Einstellung ist und wir daher auch in der Erziehung nur auf eine Einstellungsänderung hinwirken können. Aber die Autoren übersahen, dass ihr Experiment die Fähigkeitsnatur der Moral nicht widerlegt hat, da auch der zweite Schluss zutreffen könnte, nämlich dass der Test, den sie verwendeten, gar keine moralische Fähigkeit misst.

Wir haben dieses Experiment daher mit dem MKT nachgestellt, der – anders als der DIT – einen Messwert für moralische Kompetenz liefert. Auch hierbei wurden wieder Studierende untersucht, die in Gruppen gemäß der von ihnen selbst angegebenen politischen Haltung aufgeteilt wurden, und auch hier wurde der Test zweimal dargeboten, einmal zur Ermittlung der eigenen Testwerte, und dann noch einmal mit der Instruktion, beim Ausfüllen die Perspektive des politischen Gegners zu übernehmen. (Zu weiteren Details siehe Lind 2002)

Das Ergebnis unseres Experiments ist in der Abbildung auf Seite 77 dargestellt. Im Unterschied zum Emler-Experiment mit dem DIT haben die Versuchspersonen es nicht geschafft, die Testwerte in beide Richtungen zu simulieren. Es ist ihnen nicht gelungen, einen höheren C-Wert zu erhalten als in der ersten Test-Runde. Im Gegenteil, ihre Testwerte in der Simulationsaufgabe lagen immer unter denen, die sie selbst zuvor erreicht hatten. Damit ist der Schluss von Emler et al. (1983) widerlegt, das moralische Urteilsverhalten sei bloß Ausdruck einer (politischen) Einstellung und nicht einer Fähigkeit. In einem weiteren Experiment mit dem MKT hat Wasel (1993) mit einem abgewandelten Experiment unseren Befund erneut bestätigt. Er zeigte überdies, dass die Genauigkeit der Simulation hoch mit dem Niveau der eigenen moralischen Kompetenz korreliert: Demnach kann jemand mit eigener hoher Moralkompetenz diese Fähigkeit bei anderen besser erkennen als jemand mit niedriger eigener Moralkompetenz. Beide Experimente belegen, a) dass es moralische Kompetenz gibt und b) dass man sie mit dem MKT messen kann.

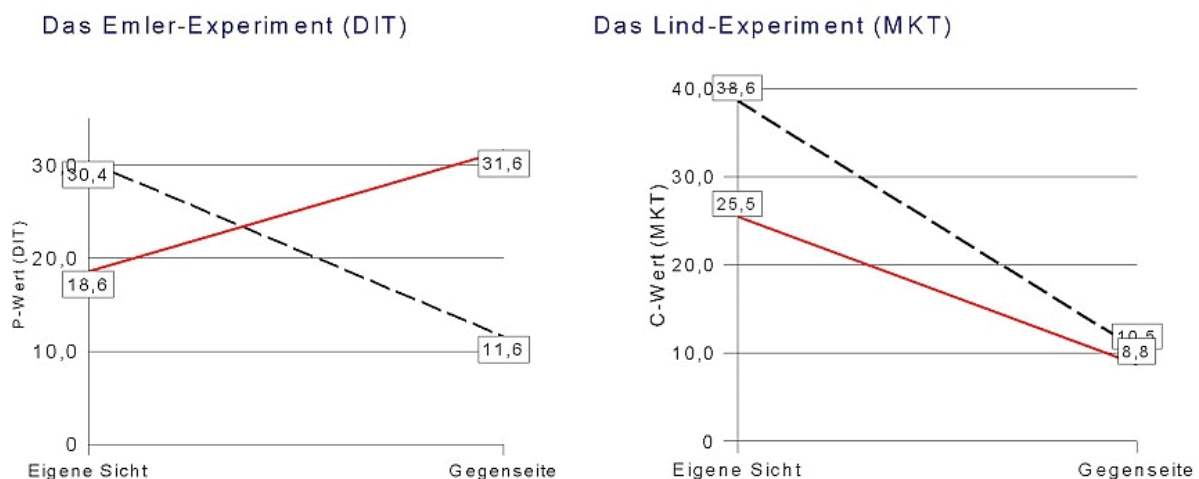


Figure 4 “Emler-Experiment”: Kann man “höhere” moralische Ideale simulieren als die eigenen? Ja. “Lind-Experiment”: Kann man “höhere” moralische Kompetenz simulieren als die eigene? Nein.