

# Sind COOL-Schüler/innen moralisch urteilsfähiger?

Ein kritisch reflektierter Antwortversuch



MAG. CHRISTOPH HELM

Wissenschaftlicher Mitarbeiter an der Abteilung für Pädagogik & Pädagogische Psychologie, Universität Linz  
christoph.helm@jku.at

Im Zentrum dieses Beitrags steht die Frage, wie mit dem Moralischen-Urteil-Test (MUT) das Konstrukt „moralische Urteilsfähigkeit“ zu messen ist sowie die Darstellung ausgewählter Ergebnisse einer Untersuchung mit diesem Instrument bei 454 Schülern/Schülerinnen berufsbildender höherer Schulen.

## 1 Forschungsfrage

Die Frage, was unter dem Bildungsbegriff zu verstehen ist, wird mindestens seit der Antike diskutiert. Ebenso lang wird behauptet, dass die Erziehung der Kinder, Jugendlichen und Erwachsenen zum autonomen, emanzipierten, mündigen, vernünftigen, ja moralischen Menschen Ziel jeder Erziehung sein müsse (vgl. KLAFKI 1993, 19). Aber inwieweit ist es gelungen, diese Ansprüche auch umzusetzen?

Mit dem bekannten Milgram-Experiment (MILGRAM 1974, 51) wurde gezeigt, dass die Gehorsamsbereitschaft (in der Moralpsychologie auch als *Autoritätsmoral* bezeichnet) von bis zu 65 % zufällig ausgewählter Probanden/Probandinnen – unter bestimmten Bedingungen – bis zum „Mord auf Befehl“ reichte. Dieses Experiment fand 1961 statt. Man sollte meinen, dass sich seither einiges – auch in der Erziehung – geändert hat. Und tatsächlich gibt es Stimmen, die vermuten, dass heute Individuen mehr Bewusstsein über die Gefahr des blinden Gehorsams gegenüber Autoritäten haben. Eine partielle Wiederholung des Experiments im Jahr 2006 zeigte jedoch, dass sich der Anteil jener, die bis zur höchsten Bestrafung bereit waren, kaum verändert hat (vgl. BURGER 2009, 1).

Schule hat die Aufgabe moralischer und sozialer Erziehung; diese ist auch im Lehrplan etwa der Handelsakademie (BMUKK 2004) zu finden. Wie aber wird diese Aufgabe (die auch eine der Eltern ist; vgl. LIND 1993, 14) erfüllt? MOHR (2006, 134, Hervorhebung im Original) kam bei einer Untersuchung von 127 Schulprogrammen deutscher Schulen zu bestärkenden Hinweisen: „Mit Abstand am stärksten ... vertreten ist die Arbeit im Bereich des *sozialen Lernens und der Werteerziehung*.“ Eine vergleichbare Studie zu österreichischen Schulen steht noch aus.

Kann schulischer Unterricht die moralische Urteilsfähigkeit von Schülern/Schülerinnen weiterentwickeln? lautete die Ausgangsfrage, der ich in einer empirischen Studie nachgehen wollte. Im Folgenden wird zunächst die Messbarkeit der abhängigen Variable erörtert, bevor ich das Forschungsdesign und einige Ergebnisse der Studie vorstelle.

## 2 Das Konstrukt „moralische Urteilsfähigkeit“

Wie kann man eine soziale Kompetenz – wie moralische Urteilsfähigkeit – überhaupt messen? Nach LEMPERS (1988, 12) wird der

Begriff „Moral“ umgangssprachlich sehr facettenreich verwendet. Da man sich auch in der moralpsychologischen und -philosophischen Wissenschaft nicht auf eine Definition einigen konnte, wird dort meist mit dem Begriff *moralischer Urteilsfähigkeit* gearbeitet, der klar zu umschreiben ist: „The capacity to make decisions and judgments which are moral (i.e., based on internal principles) and to act in accordance with such judgments.“ (KOHLEBERG 1964, 425) Dieses übereinstimmende moralische Handeln ist – wie jedes Handeln (Piaget 1976) – von einem affektiven wie auch kognitiven Aspekt getrieben. Nach LIND (2000, 40f.) sind beide Aspekte voneinander untrennbar und verlaufen parallel. LIND war einer der ersten, der versuchte, diese Parallelität zu messen, indem er eine Person dann als moralisch urteilsfähig bezeichnet, wenn sie sich bei der Beurteilung von Argumenten *konsistent an der moralischen Qualität* dieser Argumente *orientiert* (und dann erreicht sie einen hohen Wert im Moralischen-Urteil-Test; MUT). Wenn sie hingegen an ihrer „vorgefassten“ Meinung festhält, erreicht sie einen niedrigen Wert (vgl. LIND 2000, 90). Genau diese *kognitive Orientierung an der Qualität der Argumente* (z. B. in Diskussionen) versucht LIND mit dem MUT zu messen. Erfasst wird also eine Beurteilungskonsistenz; die widerspruchsfreie Argumentation hinsichtlich der Qualität der Argumente in moralisch relevanten Beurteilungssituationen (vgl. LIND 1985, 85).

## 3 Das MUT(s)-Experiment und sein 3-faktorielles Design

Bei der Messung dieser Fähigkeit im Moralischen-Urteil-Test (MUT) lesen die Probanden/Probandinnen zwei Dilemmageschichten und versetzen sich dabei in die Situation der Protagonisten/Protagonistinnen (vgl. LIND 2008/1977). Im ersten Dilemma geht es um zwei Arbeiter, die vermuten, dass die Belegschaft von der Firmenleitung illegal belauscht wird. Daher brechen sie in die Räumlichkeiten der Firmenleitung ein und stehlen die Tonbänder mit den Aufzeichnungen. Im anderen Dilemma gibt ein Arzt einer unheilbaren alten Dame Sterbehilfe, um sie von ihren Qualen zu befreien. Nach jeder dieser beiden Dilemmageschichten folgen *sechs* Argumente, die *für* die gesetzte Handlung der Akteure (Einbruch, aktive Sterbehilfe) sprechen, und dann *sechs* Argumente, die *gegen* die gesetzte Handlung sprechen. Insgesamt werden dem Probanden/der Probandin also 2 x 12 Argumente vorgelegt, die er/sie auf einer neunstufigen Skala von –4 (ich lehne das Argument völlig ab) bis +4 (ich akzeptiere das Argument völlig) bewerten soll. Jedes der Pro- und Contra-Argumente hat – der oben angeführten Definition entsprechend – unterschiedliche Qualität und entspricht genau einer (anderen) Stufe des moralischen Urteilens nach KOHLBERG (vgl. Tabelle 1).

Klassifikation der moralischen Urteile nach Entwicklungsebenen und -stufen
1. Stufe: Orientierung an Bestrafung und Gehorsam; Respekt vor überlegener Macht
2. Stufe: Richtiges Handeln ist jenes, das die Bedürfnisse des Ich und gelegentlich die der anderen instrumentell befriedigt; Orientierung an Austausch und Reziprozität
3. Stufe: Konformität mit stereotypischen Vorstellungen vom natürlichen oder Mehrheits-Verhalten
4. Stufe: Orientierung an Aufrechterhaltung von Autorität und sozialer Ordnung
5. Stufe: Anerkennung einer [...] Basis von Regeln [...] Pflicht definiert als Vertrag, allgemeine Vermeidung der Verletzung von Absichten oder Rechten anderer sowie Wille, Wohl der Mehrheit
6. Stufe: Orientierung an Gewissen oder Prinzipien ... der Entscheidung, die an logische Universalität und Konsistenz appellieren

**Tabelle 1:** Zusammenfassende Beschreibung der Stufen des moralischen Urteilens (in Anlehnung an KOHLBERG 1974, 60f.)

Der Test von LIND (1985, 82) entspricht einem 3-faktoriellen Design. Die Faktoren: *Qualität der Argumente* („Kohlberg-Stufen“), *Pro und Contra* (Meinungskonformität) und der *Dilemma-Kontext* (Arbeiter und Arzt) nehmen Einfluss auf das Antwortverhalten der Probanden/Probandinnen. Nach der Definition von LIND ist aber vor allem interessant, wie hoch der Einfluss der Stufen, d.h. der Qualität der Argumente, auf das gesamte Antwortverhalten der Probanden/Probandinnen ist. Dies lässt sich mit einer einfachen Varianzanalyse bzw. Varianzkomponentenzerlegung berechnen. Dieses Verfahren soll nun am Beispiel der Messung der Urteilskonsistenz (LIND 1985, 82ff.) überblicksartig skizziert werden. Für eine ausführliche Darstellung seien die Leser/innen auf LIND (1985) verwiesen.

#### 4 Varianzkomponentenzerlegung am Beispiel MUT

Bei der Varianzkomponentenzerlegung (einfache Varianzanalyse) wird in einem ersten Schritt das gesamte Antwortverhalten einer Person betrachtet; dazu wird die Gesamtvarianz des Antwortverhaltens ermittelt. Um diese zu berechnen, wird die jeweilige *Abweichung (der Probanden-/Probandinnenbewertung<sup>1</sup>) vom Mittelwert aller Bewertungen* berechnet und quadriert. Die durchschnittliche Größe dieser „Abweichungsquadrate“ bildet die gesuchte Gesamtvarianz.

Nach LINDs Definition interessiert aber alleine, inwieweit diese gesamte Antwortvarianz auf den Faktor „Stufe“ zurückzuführen ist. Hierzu muss zuerst der Mittelwert der Argumente sortiert je Stufe gebildet werden, damit die anderen beiden Faktoren „Dilemma“ und „Pro- und Contra“ „ausgeschaltet“ werden bzw. die Stufe die einzige „varianzgenerierende Quelle“ (BORTZ 1993, 229) darstellt. Die Abweichungen der jeweiligen *Stufenmittelwerte zum Mittelwert aller sechs Stufenmittelwerte* werden quadriert. Der Durchschnitt dieser Quadrate zeigt die Varianz an, die alleine aufgrund des Faktors „Stufe“ zustande kommt.

Setzt man „Stufenvarianz“ und „Gesamtvarianz“ in ein Verhältnis, so erhält man als rechnerisches Ergebnis dieses Experiments einen C/MUT-Wert<sup>2</sup> zwischen 0 und 100, wobei SCHILLINGER (2006, 77) folgende Kategorisierung vornimmt: „A score 1–9 is considered to be low, 10–29 medium, 30–49 high and above 50 very high.“ LIND (2008a, 18) verweist zudem darauf, dass die 100-Punkte-Skala normalerweise nur auf dem Intervall 10–40 beansprucht wird, weshalb etwa ein 10%-Unterschied bzw. eine 3-Punkte-Veränderung bereits auf einen *praktisch* bedeutsamen (signifikanten) Effekt hinweist, da diese etwa einem durchschnittlichen Zuwachs innerhalb eines Schuljahres entspricht.

Nachdem nun die Frage über das Zustandekommen der Daten zur moralischen Urteilsfähigkeit beantwortet ist, widmet sich der Beitrag im Folgenden der empirischen Antwort auf die eingangsgestellte Forschungsfrage.

#### 5 Forschungsdesign und Stichprobe der vorliegenden Untersuchung

Das Anliegen meiner Untersuchung besteht in der Beantwortung der Frage „Kann schulischer Unterricht die moralische Urteilsfähigkeit von Schülern/Schülerinnen weiterentwickeln?“ In Ermangelung von schulischen Programmen, die sich gezielt mit der Förderung „moralischer Urteilsfähigkeit“ befassen, machte ich mich auf die Suche nach Initiativen, die soziales Lernen und moralische Erziehung in einem weiteren Sinn zu ihren Zielen zählen – in der Hoffnung, dadurch das Phänomen „Moralische Urteilsfähigkeit“ besser zu verstehen und erste Erfahrungen mit seiner Messung machen zu können.

Die wohl größte Schulentwicklungsbewegung der österreichischen Sekundarstufe II ist die Initiative COoperatives Offenes Lernen (COOL) mit mittlerweile über 30 Impulsschulen (vgl. COOLTRAINERS 2010). Wiewohl COOL nun kein spezifisches und gezieltes Programm zur Moralförderung durch Schule darstellt und auch diese nicht unter seinen vordersten Zielen genannt wird, so schien es mir trotzdem interessant, etwaige vorhandene (oder fehlende) moralförderliche Effekte dieses Schultyps einmal explorativ zu untersuchen.

Die Förderung moralischer Urteilsfähigkeit von Schülern/Schülerinnen gehört nicht nur zu den Bildungszielen *österreichischer Lehrpläne* (vgl. BMUKK 2010); in der Betonung von *kooperativem, sozialem Lernen* durch COOL (vgl. WITTEW/NEUHAUSER 2008, 25f.) scheint „moralförderliches“ Potenzial zu stecken. In einer konzeptuellen Analyse (Helm 2009; im Erscheinen) habe ich die drei Konzepte „COOL/Daltonplan“ (vgl. PARKHURST 1922; WITTEW/NEUHAUSER 2008), „Just Community“ (vgl. KOHLBERG 1986) und die „moralförderlichen Entwicklungsbedingungen“ (vgl. LEMPERS 1988) verglichen. Als Ergebnis kann festgehalten werden, dass COOL aufgrund von Elementen wie Klassenrat, Lehrer/-innenkooperation, Einbindung der Eltern, freie, selbsttätige Lernphasen, Gruppen- und Teamarbeiten, fächerübergreifender, vernetzter Unterricht ... besser als „herkömmlicher“ Unterricht dazu geeignet sein müsste, die soziale Kompetenz „moralische Urteilsfähigkeit“ zu fördern. Sowohl die didaktische Konzeption von COOL wie auch seine *Grundhaltung der „positiven pädagogischen Zumutung“* (vgl. OSER/ALTHOF 1997, 353), dass Schüler/innen Verantwortung übernehmen können und ihnen dies auch gewährt wird, müssten dies ermöglichen.

Aufgrund dieser Überlegungen habe ich meine Ausgangsfrage „Kann schulischer Unterricht die moralische Urteilsfähigkeit von Schülern/Schülerinnen weiterentwickeln?“ hier zur Forschungsfrage „Welchen Beitrag leistet COOL zur Förderung moralischer Urteilsfähigkeit?“ konkretisiert.

Diese Frage kann im strengen Sinn nur mittels einer Längsschnittstudie beantwortet werden. Aus Ressourcengründen muss-

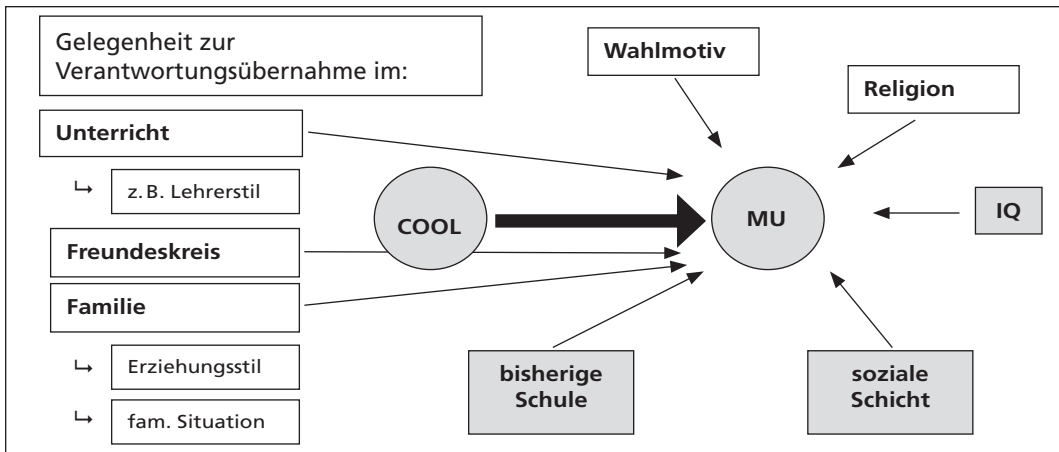


Abbildung 1: Zusammenhänge untersuchter Variablen

te ich mich hier mit einigen Querschnittsvergleichen begnügen und hoffte damit erste Indizien zu finden, die die Ergebnisse meiner oben erwähnten konzeptuellen Analyse auch empirisch stützen können.

Die *Probanden/Probandinnen* der Untersuchung sind Schüler/innen der Sekundarstufe II. Nach Kontakt mit dem COOL-Impulszentrum wurden zwei Handelsakademien und eine höhere Lehranstalt für wirtschaftliche Berufe ausgewählt, die das COOL-Konzept bereits in allen fünf Jahrgängen der höheren Schule umsetzen und von denen bekannt ist, dass das COOL-Konzept bereits „gut läuft“. Die Kontrollgruppen setzten sich aus, je Schule unterschiedlichen, Ausbildungszweigen (Management und Controlling, Informationstechnologie ...) zusammen. In den Monaten Jänner und Februar 2009 wurde nach einer Vortestung der verwendeten Instrumente die Untersuchung durchgeführt. Insgesamt umfassen die untersuchten Schulklassen 524 Schüler/innen, wobei letztlich 454 (rund 87%) an der Befragung teilnahmen. Folgende Abschnitte wurden von den Probanden/Probandinnen bearbeitet: Moralischer-Urteil-Test (MUT/s, /s = Schülerversion)<sup>3</sup>, Skalen zur sozialen Beziehungen in Familie, Freundeskreis und Schule, zur Religion, zum COoperativen Offenen Lernen, zur sozialen Schicht sowie zu biografischen Daten<sup>4</sup>. Damit wurden die in Abbildung 1 dargestellten Variablen erhoben, wobei im Folgenden nur auf die grau hervorgehobenen eingegangen wird, da – in meinen Daten – nur diese Einfluss ausüben.

## 6 Ausgewählte Ergebnisse der Untersuchung

In der vorliegenden Untersuchung wurde an 454 Schülern/Schülerinnen das MUT/s-Experiment durchgeführt und die durchschnittlichen C-Werte errechnet. Abbildung 2 (unten) zeigt die Verteilung kategorisiert nach SCHILLINGER (2006, 77) zwischen Versuchs- und Kontrollgruppe an. Die Hypothese, dass COOL-Schüler/innen besser abschneiden, kann mit vorliegenden Daten nicht bestätigt werden. Zwar ist ihr Anteil im Bereich „medium“ deutlich höher, jedoch zeigt ein direkter Mittelwertvergleich über alle Kategorien hinweg keine statistisch und praktisch signifikanten Unterschiede (siehe Abbildung 2). Etwas Aufmerksamkeit sollte in Abbildung 2 auch darauf gerichtet werden, dass der Anteil der untersten Kategorie „low“ in der gesamten Stichprobe doch hoch ist.

Wenn nicht die „Unterrichtsform“ COOL oder Nicht-COOL<sup>5</sup> den MUT/s-Wert beeinflusst, welche der überprüften Variablen sind es dann? Die Antwort findet sich zum Teil in kognitiven Fähigkeiten – die in dieser Untersuchung durch die Noten abgebildet werden<sup>6</sup> – bzw. in der sozialen Schicht. Statistisch gesehen korreliert der „Anstieg“ des Notengrades (hin zur besseren Note) über die sechs erhobenen Fächer hinweg mit .17 nur schwach, jedoch auf einem Niveau von 0,01 hoch signifikant. Die praktische Bedeutsamkeit der *Note* dagegen ist als relativ stark zu bewerten, wie aus Abbildung 3 deutlich wird. Die Spannweite der MUT/s-Werte reicht von etwa 12 Punkten bei Note 5 bis etwa 21 Punkte

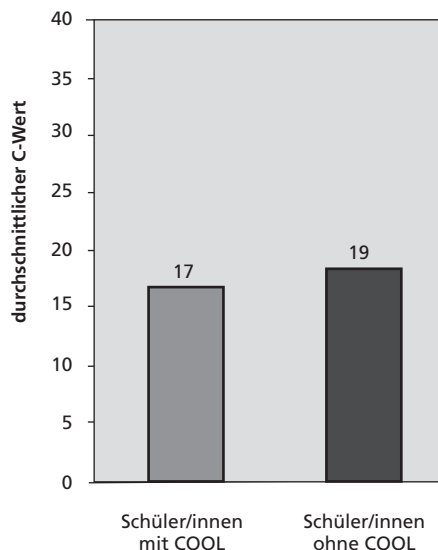
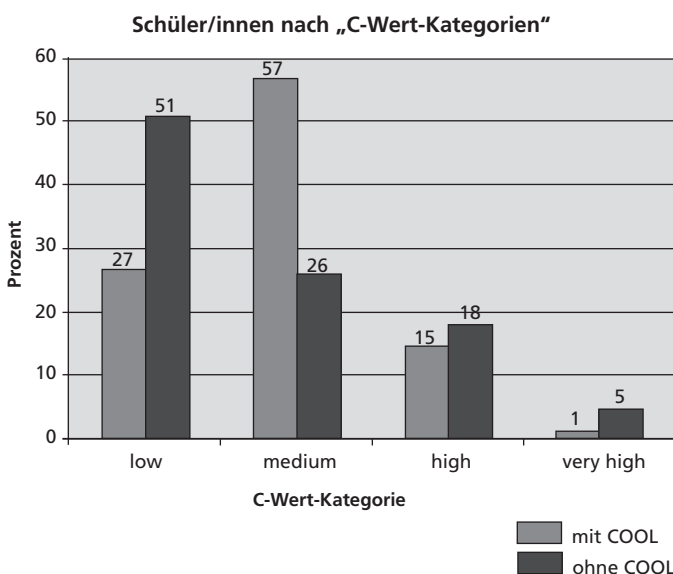
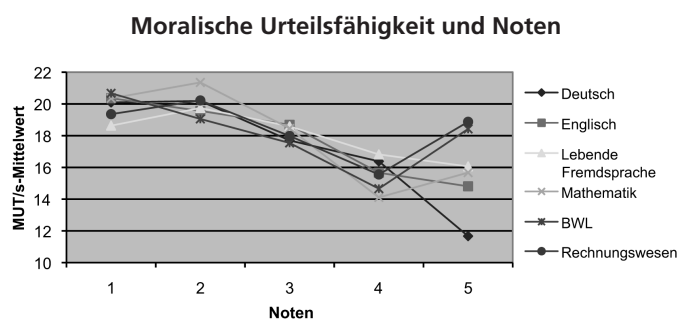


Abbildung 2: Vergleich der Versuchs- und Kontrollgruppe nach Aufteilung je „C-Wert-Kategorie“ (links) sowie nach dem C-Mittelwert (rechts).

bei Note 1. Etwas verschärft formuliert, hat also die Schulnote auf etwa ein Drittel der üblicherweise beanspruchten MUT(/s)-Skala Einfluss, unterstellt man auf Basis theoretischer Überlegungen und empirischer Befunde einen kausalen Zusammenhang. Mein Ergebnis steht dabei in einer Spannung zu einer aktuellen Studie von HETTINGER (2009), die allerdings kognitive Fähigkeiten über die Intelligenz erhob. HETTINGER (2009, 60) musste die Hypothese, dass sich überdurchschnittlich und hoch Begabte durch eine höhere moralische Urteilsfähigkeit auszeichnen als durchschnittlich Begabte, verwerfen.



**Abbildung 3:** Moralische Urteilsfähigkeit (C- bzw. MUT/s-Mittelwerte) je Fächernote

Einen praktisch und statistisch signifikanten Einfluss auf die moralische Urteilsfähigkeit zeigt auch der *soziale Hintergrund der Schüler/innen*. Den Ergebnissen meiner Untersuchung zufolge steigt der C-Werte mit der sozialen Schicht, wobei mit der gehobenen Schicht ein Deckeneffekt eintritt. Der praktische Unterschied liegt hier bei etwa 5,5 MUT/s-Punkten, was laut LIND (1993, 181) der durchschnittlichen Entwicklung innerhalb ein bis zwei normaler Schuljahre entspricht. Ähnliche Ergebnisse finden sich bei WEYERS (2004, 185).

soziale Schicht	Mittelwert MUT/s	N	%
hohe	18,09	15	3,33
gehobene	20,68	149	33,11
mittlere	17,35	248	55,11
niedrige	14,53	38	8,44
gesamt	18,24	450	100,00

**Tabelle 2:** Durchschnittliche MUT/s-Werte je sozialer Schicht

Die hier sehr verkürzte Darstellung<sup>7</sup> der empirischen Ergebnisse schließt mit den Ergebnissen des *Unterschlagungsexperiments*, das am Ende der Untersuchung durchgeführt wurde. Mit diesem wurde der Frage nachgegangen, ob und inwiefern die erhobene moralische Urteilsfähigkeit auch mit realem Handeln der Schüler/innen zusammenhängt. Ich vermutete, dass Schüler/innen, die sich an eine Vereinbarung halten, über durchschnittlich höhere C-Werte verfügen. In Anlehnung an KREBS/ROSENWALD (1977) wurde mit den Schülern/Schülerinnen direkt nach der Testung eine Art Vertrag geschlossen, in dem sie sich verpflichteten, binnen zwei Wochen Onlinefragen zum Thema Moral zu beantworten. Um den Ernstcharakter dieser Vereinbarung deutlich zu machen, wurde mehrmals darauf hingewiesen, dass die fristgerechte Beantwortung dieser Fragen für die wissenschaftliche Untersuchung von hoher Bedeutung sei. Die Rücklaufquote betrug etwa 30% (31% bei COOL- und 25% bei NonCOOL-Schülern/Schülerinnen). Die durchschnittliche moralische Urteilsfähigkeit jener, die sich an die

Vereinbarung hielten, lag etwa zwei Punkte höher als die derjenigen, die sich nicht an die Vereinbarung hielten. Dies ist jedoch statistisch und praktisch nicht signifikant.

Unterschlagung	N	Mittelwert MUT/s
fristgerechte Antwort	129	19,90
Antwort unterschlagen	318	17,49

**Tabelle 3:** MUT/s-Mittelwertunterschiede der Unterschlagungs- und Kooperationsgruppe

Zusammenfassend kann festgehalten werden, dass die vorliegende Querschnittstudie zu folgenden Ergebnissen kam:

- 1 Eine Analyse des Konzepts „Just Community“ sowie der moralfördernden Entwicklungsbedingungen nach Lempert (1998) zeigte, dass COOL aus konzeptueller Sicht sehr gut zur Förderung/Entwicklung sozialer Kompetenzen (respektive moralischer Urteilsfähigkeit) geeignet ist.
- 2 Die empirischen Ergebnisse der vorliegenden Studie weisen jedoch in eine andere Richtung. Ihr zufolge scheinen COOL-Schüler/innen hinsichtlich der moralischen Urteilsfähigkeit nicht mehr von COOL zu profitieren als ihre Kollegen/Kolleginnen der Kontrollgruppe vom tradierten Unterricht. Jedoch sind diese Ergebnisse nur vorläufige. Eine Längsschnittstudie ist abzuwarten.
- 3 Zusammenhänge zwischen COOL und moralischer Urteilsfähigkeit konnten in einer explorativen Analyse nicht gefunden werden. Hingegen zeigte sich, dass die Variablen „soziale Schicht“ und „Schulnoten“ (statistisch und praktisch) signifikanten, positiven Einfluss auf den MUT/s-Wert ausüben.

## 7 Methodenkritik

Fragen zum Design einer Untersuchung sind immer *auch* Ressourcenfragen. Dennoch muss der Frage nachgegangen werden, ob die angewandten Methoden gegenstandsangemessen sind (vgl. FLICK 1995, 13ff.). Ziel der empirischen Untersuchung war es, die Ergebnisse der konzeptuellen Analyse bzw. die daraus abgeleitete Hypothese mit Daten zu belegen. In diesem Sinne stellt der Forschungsprozess eine deduktive Vorgehensweise dar. Für diese Hypothesentestung sind quantitative bzw. experimentelle Methoden besser geeignet als qualitative Verfahren (vgl. POPPER 1984/1934). Weiters lässt sich die Vorgehensweise u. a. aufgrund des hohen Stichprobenumfangs rechtfertigen.

Kritisch zu hinterfragen ist jedoch, inwiefern es gerechtfertigt ist, aus der konzeptuellen Analyse abzuleiten, dass COOL-Schüler/innen stärker bezüglich der moralischen Urteilsfähigkeit gefördert werden? Hierzu hätte sicherlich eine qualitative Analyse, z.B. durch Unterrichtsbeobachtung und Betroffeneninterviews mit Lehrkräften und Schülern/Schülerinnen, Aufschluss darüber gegeben, ob eine derartige Annahme, nicht nur aus der Sicht der Konzeptpapiere, sondern auch aus Sicht der Schulpraxis, gerechtfertigt ist.

Würde eine derartige qualitative Analyse zu einem positiven Ergebnis führen, müsste man weiters noch kritisch hinterfragen, ob mit dem eingesetzten MUT(/s) auch das gemessen wird, was im Unterricht gelehrt wird bzw. was COoperatives Offenes Lernen stärker fördert, als „herkömmlicher“ Unterricht. D.h. auch die Gegenstandsangemessenheit des eingesetzten Tests muss hinterfragt werden. Dies war zentrales Anliegen des Beitrags: Indem die Idee des MUT(/s) und seine Auswertung dargestellt wurden, wurde aufgezeigt, wie die empirischen Daten zustande kommen. Auf

Basis dieser Erkenntnis sollte es den Lesern/Leserinnen ermöglicht werden, die vom mir dargestellten Ergebnisse im Lichte ihres Zustandekommens zu beurteilen. Zum MUT(/s) selbst lässt sich kritisch anmerken, dass der Zusammenhang zur Note und zur sozialen Schicht darauf hindeutet, dass eine stark kognitive Fähigkeit gemessen wird, was ja – neben der affektiven Komponente – auch seine Intention ist (vgl. LIND 1982, 4). Gleichzeitig muss mit HETTINGER (2009, 50) aber auch darauf verwiesen werden, dass dabei möglicherweise bildungsferne bzw. sozioökonomisch „niedrige“ Schichten benachteiligt werden. Insofern müsste man den oben dargestellten Einfluss der sozialen Schicht als einen vom Testinstrument transintentionalen Effekt interpretieren. Zwar soll nach LIND (vgl. 2000, 89) der MUT die moralische Urteilsfähigkeit als eine kognitive und affektive Fähigkeit messen, aber sicherlich keine, die etwa auf die Lesekompetenz zurückzuführen ist.

Das Ergebnis der vorliegenden Untersuchung, dass COOL-Schüler/innen im Schnitt bezüglich der erhobenen Kompetenz nicht besser, aber auch nicht schlechter sind, als Schüler/innen anderer HAK/HLW-Formen muss im Lichte dieser Einschränkungen interpretiert werden. Und selbst dann muss eine weitere Relativierung vorgenommen werden, nämlich dass offenbar der Zusammenhang zwischen getroffenem moralischem Urteil und der darauf gesetzten Handlung nicht linear verläuft, sondern auch (?) von anderen Umständen beeinflusst wird (vgl. OSER/ALTHOF 1998, 225). Auch LIND verweist darauf, dass mit dem MUT keinerlei Anspruch darauf besteht, einen „bestimmten Verhaltensakt eines Menschen vorherzusagen“ (LIND 1993, 24, Hervorhebung im Original), auch wenn beides miteinander korreliert (vgl. LIND 2007a). ✘

## Anmerkungen

- 1 Jedes Item wird auf einer Zustimmungsskala von -4 bis +4 beurteilt.
- 2 „C“ steht für Kompetenz-Wert und soll ausdrücken, dass es sich hier um eine kognitive Fähigkeit handelt (vgl. LIND 2007).
- 3 Zwar sollte nach LIND (2009) die Originalversion des MUT durchaus für die vorliegende Stichprobe geeignet sein, jedoch zeigte die Voruntersuchung, dass die Schüler/innen mit sprachlichen Problemen zu kämpfen hatten. Aus diesem Grund wurden einige wenige Wörter adaptiert, woraus die Bezeichnung MUTs und eine Validitätsprüfung resultierte (HELM 2009, 61ff.). Daher können die Ergebnisse der vorliegenden Studie nur vor diesem Hintergrund mit den Ergebnissen von LIND verglichen werden.
- 4 Weiters wurde ein Unterschlagungsexperiment durchgeführt (vgl. HELM 2009, 132ff.).
- 5 Ein methodisches Problem der Studie besteht darin, dass COOL in jeder Schule unterschiedlich intensiv betrieben wird und daher etwas anders ausgeprägt ist. Zum anderen können und sollen auch in NonCOOL-Klassen die moralische Urteilsfähigkeit fördernde Elemente (etwa vermehrter Einsatz von Sozialformen im Unterricht) vorkommen. Um dieser „Ausprägungsproblematik“ gerecht zu werden, wurden adäquate Maßnahmen gesetzt, die in HELM (2009, 96ff.) nachzulesen sind.
- 6 Aus methodischen und zeitlichen Gründen war der Einsatz eines Intelligenztests nicht möglich, weshalb in der vorliegenden Untersuchung Schulnoten als Indikator für kognitive Fähigkeiten herangezogen wurden. Zwar verweist bereits z. B. INGENKAMP (1975) darauf, dass Ziffernnoten ihre Ansprüche an Objektivität, Reliabilität und Validität nicht einlösen können. Dennoch empfiehlt LIND (2009a) Schulnoten zu erheben, da für sie eine höhere prognostische Validität als bei Intelligenztests spricht. Es wurden jene Fächer ausgewählt, die in Summe über 50 % der Stundentafel ausmachen, um so einen möglichst großen Anteil der in der Schule verbrachten Zeit abzudecken.
- 7 Für eine genauere Analyse der empirischen Daten vgl. HELM (2009; sowie im Erscheinen).

## Literatur

BMUKK (2004): Lehrplan der Handelsakademie. BGBl. II – 19. Juli 2004 – Nr. 291. Online: [http://www.bmukk.gv.at/medienpool/11701/vo\\_aend\\_lp\\_bak\\_hasch\\_anl1.pdf](http://www.bmukk.gv.at/medienpool/11701/vo_aend_lp_bak_hasch_anl1.pdf) (01. 08. 2009).

- BMUKK (2010): Soziales Lernen. Online: [http://www.bmukk.gv.at/schulen/unterricht/ba/soziales\\_lernen.xml](http://www.bmukk.gv.at/schulen/unterricht/ba/soziales_lernen.xml) (08. 06. 2010).
- BORTZ, J. (1993): Statistik für Sozialwissenschaftler. Berlin.
- BURGER, J.M. (2009): Replicating Milgram. Would People Still Obey Today? In: *American Psychologist*, 64(1), 1–11.
- COOLTRAINERS (2010): Netzwerkpartner. Online: <http://www.cooltrainers.at/index.php?id=183> (13. 04.2010).
- FLICK, U. (1995): Qualitative Forschung. Reinbek bei Hamburg.
- HELM, C. (2009): Wie moralisch ist COOL? Welchen Beitrag leistet COoperatives Offenes Lernen zur Entwicklung moralischer Urteilsfähigkeit? Diplomarbeit. Abteilung für Pädagogik und Pädagogische Psychologie, Linz: Johannes Kepler Universität.
- HELM, C. (im Erscheinen): Welchen Beitrag leistet COoperatives Offenes Lernen zur Entwicklung moralischer Urteilsfähigkeit? Erscheint im Tagungsband zur Österreichischen Konferenz für Berufsbildungsforschung.
- HETTINGER, G. (2009): Vergleich von moralischer Urteilskompetenz und Werthaltungen bei durchschnittlich begabten und weit überdurchschnittlich/hoch begabten Jugendlichen. In: KERNER, H.-J. (Hrsg.): *Tübinger Schriften und Materialien zur Kriminologie*. Tübingen.
- KLAFKI, W. (1993): *Neue Studien zur Bildungstheorie und Didaktik: zeitgemäße Allgemeinbildung und kritisch-konstruktive Didaktik*. Basel.
- KOHLBERG, L. (1964): *Development of moral character and moral ideology*. In: HOFFMAN, M.L./HOFFMAN, L.W. (Eds.): *Review of child development research*. New York, 381–431.
- KOHLBERG, L. (1974): *Zur kognitiven Entwicklung des Kindes*. Frankfurt/Main.
- KOHLBERG, L. (1986): *Der „Just Community“-Ansatz der Moralerziehung in Theorie und Praxis*. In: OSER, F./FATKE, R./HÖFFE, O. (Hrsg.): *Transformation und Entwicklung*. Frankfurt/Main, 21–55.
- KREBS, D./ROSENWALD, A. (1977): *Moral reasoning and moral behavior in conventional adults*. In: *Merrill-Palmer Quarterly*, 23, 79–84.
- LEMPERT, W. (1988): *Moralisches Denken. Seine Entwicklung jenseits der Kindheit und seine Beeinflussbarkeit in der Sekundarstufe II*. Essen.
- LIND, G. (1982): *Der „Moralisches-Urteil-Test“ (m-u-t) in der Kritik*. Forschungsbericht. Universität Konstanz.
- LIND, G. (1985): *Inhalt und Struktur des moralischen Urteilens*. Dissertation. Universität Konstanz.
- LIND, G. (1993): *Moral und Bildung. Zur Kritik von Kohlbergs Theorie der moralisch-kognitiven Entwicklung*. Heidelberg.
- LIND, G. (2004): *The Meaning and Measurement of Moral Judgment Competence: A Dual-Aspect Model*. In: FASKO, D./WILLIS, W. (Eds.): *Contemporary Philosophical and Psychological Perspectives on Moral Development and Education*. Cresskill NJ, 185–220.
- LIND, G. (2000): *Ist Moral lehrbar?* Berlin.
- LIND, G. (2007): *Kurze Beschreibung des Moralischen Urteil-Tests (MUT)*. Online: <http://www.uni-konstanz.de/lag-moral/mut/MJT-d-kurz.htm> (13. 01. 2009).
- LIND, G. (2007a): *Die Bedeutung von moralischer Urteilsfähigkeit. Rückblick auf 30 Jahre moralpsychologischer Forschung*. Online: [www.uni-konstanz.de/lag-moral.../2\\_verhaltens-relevance\\_d.pdf](http://www.uni-konstanz.de/lag-moral.../2_verhaltens-relevance_d.pdf) (13. 12. 2008).
- LIND, G. (2008/1977): *The Moral Judgment Test (MJT) – German version*. (Um Testunterlagen kann unter [georg.lind@uni-konstanz.de](mailto:georg.lind@uni-konstanz.de) direkt bei Georg Lind ange-sucht werden.)
- LIND, G. (2008a): *Effektstärken: Statistische versus praktische und theoretische Bedeutsamkeit*. Forschungsbericht. Konstanz.
- LIND, G. (2009): *Moral Judgment Test. Frequently Asked Questions*. Online: <http://www.uni-konstanz.de/lag-moral/mut/mjt-engl.htm> (13. 04. 2010).
- LIND, G. (2009a): *Der „Moralisches Urteil-Test“ in Zusammenhang mit Noten*. E-Mail: [georg.lind@uni-konstanz.de](mailto:georg.lind@uni-konstanz.de) (13. 01. 2009).
- MILGRAM, S. (1974): *Das Milgram-Experiment*. Hamburg.
- MOHR, I. (2006): *Analyse von Schulprogrammen*. München.
- OSER, F./ALTHOF, W. (1997): *Moralische Selbstbestimmung: Modelle der Entwicklung und Erziehung im Wertebereich*. Ein Lehrbuch. Stuttgart.
- PARKHURST, H. (1922): *Education On The Dalton Plan*. New York.
- POPPER, K.R. (1984/1934): *Logik der Forschung*. Tübingen.
- SCHILLINGER, M. (2006): *Learning environment and moral development: How university education fosters moral judgment competence in Brazil and two German-speaking countries*. Aachen.
- WEYERS, S. (2004): *Moral und Delinquenz. Moralische Entwicklung und Sozialisation straffälliger Jugendlicher*. München.
- WITTWER, H./NEUHAUSER, G. (2008): *COOL Cooperatives offenes Lernen: eine Initiative für mehr Selbständigkeit, Eigenverantwortung und Kooperation an unseren Schulen*. Beibehft zum Film. Wien.