

Zur Messung sozialer Motivation in der BIJU-Studie¹

Georg Lind²
Universität Konstanz

¹ Diese Stellungnahme wurde im Auftrag und mit Unterstützung der Gewerkschaft Erziehung und Wissenschaft, GEW, erstellt. Die Verantwortung für die hierin zum Ausdruck gebrachten Meinungen liegt jedoch allein beim Autor.
Weitere Artikel zu der BIJU-Studie hat die Gemeinnützige Gesellschaft Gesamtschule (GGG) von Nordrhein-Westfalen auf ihrer Web-Seite veröffentlicht (<http://www.GGG-NRW.de/>)

² Anschrift des Verfassers: Prof. Dr. Georg Lind, Sozialwissenschaftliche Fakultät, D-78457 Konstanz. E-mail: Georg.Lind@uni-konstanz.de ; URL: <http://www.uni-konstanz.de/ag-moral/>

Zusammenfassung

Die Studie "Bildungsverläufe und psychosoziale Entwicklung im Jugendalter (BIJU)" ist - nach Aussage des Projektleiters Professor Dr. Jürgen Baumert und seines Mitarbeiters Olaf Köller - so angelegt, dass sie nur die *Beschreibung* ("Deskription") eines Zustands zulässt, aber keine Erklärung oder Zuschreibung von Ursachen und Effekten (S. 13 und 18). Gleichwohl erwarten sich die Autoren davon vieles: "Nützliches Orientierungswissen" und "Klarheit über die Ertragslage" des Bildungssystems, insbesondere der verschiedenen Schulformen. Aus diesem offenkundigen Widerspruch in den Absichten der Autoren der BIJU-Studie resultieren, wie die eingehende Analyse zeigt, viele Probleme. Die Anlage der Studie und ihre Daten stellen kaum eine tragfähige Basis für die zum Teil weitreichenden Schlussfolgerungen der Autoren und vieler Rezipienten in der Öffentlichkeit dar, zum Beispiel zu den Auswirkungen der Schulform auf die Entwicklung psychosozialer Fähigkeiten. Die Analyse der BIJU-Studie zeigt im einzelnen:

- Die Absicht der Autoren, Fähigkeiten im psychosozialen Bereich zu erfassen, ist in dem Messinstrument der BIJU-Studie nur ungenügend, wenn überhaupt repräsentiert. Prosoziale Motive sind keine Fähigkeiten. Motive eignen sich nicht zur Evaluation von Bildungsprozessen oder Schulformen. Eine solche Evaluation ist nicht vereinbar mit dem Selbstverständnis von Bildung in einer demokratischen Gesellschaft.
- Die Tatsache, dass die Nennung altruistischer Motive mit bestimmten Kompetenzen korreliert, ist kein Beleg dafür, dass die von den Befragten berichteten Motive selbst auch Kompetenzen darstellen.
- Selbst wenn man die Messung prosozialer Motivation für sich genommen als ein sinnvolles Unterfangen ansieht (wenn man sie also nicht als Indikator für psychosoziale Fähigkeiten nimmt und nicht zur Evaluation von Schulformen heranzieht), weist die Messmethode der BIJU-Studie schwere Validitätsmängel auf, die eine eindeutige Interpretation der Befunde erschwert, wenn nicht gar unmöglich macht.
- Andere, bereits seit langem vorhandene Messansätze, die Fähigkeiten im psychosozialen Bereich (wie "moralische Urteilskraft") direkt und valide erfassen, wurden offenbar nicht in Betracht gezogen.
- Schließlich wurden in der BIJU-Studie die Vorteile der Längsschnittforschung, mit der in der Regel hohe Kosten verbunden sind, bislang nicht genutzt, um alternative Erklärungen für unterschiedliche psychosoziale Verän-

derungen (wie z.B. Anfangsunterschiede) empirisch zu testen. Es fehlt eine kritische Analyse der für Längsschnittstudien typischen Verzerrungseffekte (wie Selektion, Ausfälle und soziale Erwünschtheit). So ist nicht klar, ob die Unterschiede der "prosozialen Motivation" zwischen den Schulformen, wie die Autoren meinen, wirklich schulformbedingt ist oder eher durch andere Ursachen, wie z.B. Selektionseffekte oder soziale Erwünschtheit.

Vorbemerkungen

Im folgenden werden anhand von Zitaten der Leiter und Mitarbeiter der BIJU-Studie ihre Ziele, Methoden und Ergebnisse kurz dargestellt. Diese werden dann jeweils kommentiert.

Der Schwerpunkt meiner Analyse liegt auf der Frage, ob die von der BIJU-Studie gelieferten Daten eine tragfähige, valide Basis für die Interpretationen und Befunddarstellungen der Autoren darstellen. Ich werde untersuchen, inwieweit die Anlage der BIJU-Studie dem bekundeten Untersuchungsinteresse entspricht, inwieweit die Beobachtungsabsichten durch die gewählten Messinstrumente abgedeckt werden und schließlich, welche Vorkehrungen getroffen wurden, um die Validität der Messungen bzw. Beobachtungen sicherzustellen, indem alternative Interpretationsmöglichkeiten empirisch überprüft werden.

Meine Analyse ist durchweg von der Annahme getragen, dass eine solch aufwendige, aus öffentlichen Mitteln finanzierte Studie der wissenschaftlichen Erkenntnisgewinnung und letzten Endes auch der Verbesserung unseres Bildungssystems dienen muß, und nicht bloß der Befriedigung einer diffusen Neugier. Auch wenn die Autoren den rein "deskriptiven" Charakter ihrer Studie betonen, greifen sie mit ihren Publikationen und Vorträgen massiv in wissenschaftliche und bildungspolitische Diskussionen ein, was die Behauptung, mit der BIJU-Studie solle nur beschrieben werden, wenig überzeugend erscheinen läßt.³

³ Die moderne Wissenschaftstheorie zeigt, dass eine "reine Beschreibung", ohne Bezug zu Theorien und Erkenntnisinteressen nicht möglich ist. Die Auswahl der Entwicklungsziele und -faktoren und der Befragtengruppen erfordert Entscheidungen, die immer interessen- und theoriebestimmt sind. Selbst so einfach erscheinende Dinge wie die Längenmessung stützen sich, wie Popper (1968) schreibt, auf Theorien: "There is no measurement without a theory and no operation which can be satisfactorily described in non-theoretical terms. The attempts to do so are always circular; for example, the description of the measurement of length needs a (rudimentary) theory of heat and temperature-measurement; but these, in turn, involve measurements of length." (S. 62)

Die BIJU-Studie

Untersuchungsinteresse

Die Studie "Bildungsverläufe und psychosoziale Entwicklung im Jugendalter (BIJU)" ist - nach Aussage des Projektleiters Professor Dr. Jürgen Baumert und seines Mitarbeiters Olaf Köller - so angelegt, dass sie nur die *Beschreibung* ("Deskription") eines Zustands zulässt, aber keine Erklärung oder Zuschreibung von Ursachen und Effekten (S. 13 und 18). Eine ähnliche, rein beschreibende Absicht lag auch schon der TIMS-Studie zugrunde. Gleichwohl erwarten sich die Autoren von solchen reinen Beschreibungen vieles: "Nützliches Orientierungswissen" und "Klarheit über die Ertragslage" des Bildungssystems, insbesondere, wie ihre Ergebnisdarstellung zeigt, der verschiedenen Schulformen.

Kommentar

- Offenkundig sind die Untersuchungsinteressen der Autoren der BIJU-Studie untereinander nicht kompatibel. Orientierungswissen und Klarheit über die Ertragslage verschiedener Schulformen und anderer Aspekte des Bildungssystems könnten nur gewonnen werden, wenn die Studie auf dieses Untersuchungsinteresse hin angelegt worden wäre. Gerade dies ist aber laut Autoren nicht der Fall. Sie ist absichtlich nur auf bloße Deskription angelegt. Damit ist aber keine Klärung von Ursachen und Effekten möglich.
- Alles deutet darauf hin, dass von den beiden Untersuchungsinteressen das letztlich entscheidende die Klärung von Effekten ist. Reine Deskription hätte wohl kaum die Durchführung einer so umfangreichen und teuren Studie wie BIJU gerechtfertigt, die ausschließlich durch öffentliche Mittel aus dem Bildungshaushalt finanziert wird.
- Gegen das angebliche Interesse einer reinen Deskription spricht auch die Tatsache, dass die Autoren keineswegs bei der Beschreibung stehen bleiben, sondern alle Befunde zur Erklärung von Kausal- und Effektzuschreibungen benutzen. Im Mittelpunkt der Befunddarstellung steht fast ausschließlich die Frage, inwieweit die Schulform (Gesamtschule versus andere Schulformen) für Unterschiede hinsichtlich der psychosozialen Entwicklung von Jugendlichen verantwortlich gemacht werden können.

Fazit: Die BIJU-Studie wurde durchgeführt, um wichtige erziehungswissenschaftliche Fragen und bildungspolitische Kontroversen zu klären. Sie ist jedoch dafür nicht geeignet, da sie rein “deskriptiv” ist, also nicht so angelegt, dass sie diesen Zweck erfüllen kann.

Messintention

Mit der BIJU-Studie soll laut Autoren “die Entwicklung von Jugendlichen und jungen Erwachsenen im kognitiven und psychosozialen Bereich in Abhängigkeit von unterschiedlichen familiären und institutionellen Entwicklungsbedingungen untersucht” werden (Baumert & Köller, 1998, S. 13).

Die Studie konzentriert sich besonders auf die “*sozialen und emotionalen Fähigkeiten*”, die ebenso wie die “kognitiven Fähigkeiten” Zielsetzung schulischer Bildungsprozesse seien (S. 15, meine Hervorhebung).

Ein Kernbereich der psychosozialen Fähigkeiten, die mit der BIJU-Studie erfasst werden sollen, sind “*psychosoziale Motive* und die Entwicklung *moralischer Urteilskraft*” (S. 16 und 17).

Bei den psychosozialen Motiven (oder moralischen Urteilskraft) werden drei Typen unterschieden: *Egoismus*, *Konformität* und *Altruismus*.

Die Autoren der BIJU-Studie gehen davon aus, a) dass es sich bei diesen Typen um eine Charakterisierung von unterschiedlichen *Fähigkeiten* handelt und b) dass diese Typen (Egoismus, Konformität, Altruismus) eine *Entwicklungsskala* bilden, das heißt, dass Egoismus eine niedrige psychosoziale Fähigkeit darstellt als Konformismus und Altruismus, und Altruismus die höchste (S. 16, 17, 18; siehe auch oben).⁴

Kommentar

- Es ist nicht erkennbar, wo der Anspruch der BIJU-Studie, Fähigkeiten im psychosozialen Bereich zu erfassen, eingelöst wird. Im ganzen Instrumentarium der BIJU-Studie finden sich nur relativ wenige Hinweise auf Fragen und Aufgaben, die Fähigkeiten zu messen beabsichtigen. Unter den sieben Sektionen des Erhebungsinstrumentes befinden sich zwei, die Hinweise auf die Erhebung

⁴ In anderen Veröffentlichungen werden diese Projektziele noch ergänzt um die Begriffe “Persönlichkeitsentwicklung” (MPI Zwischenbericht 1994, S. 7) und “soziale Kompetenzen” (BIJU Erhebungsinstrument, ohne Datum).

von Fähigkeiten oder Kenntnissen enthalten: die Sektion "Soziale Kompetenz, soziale Integration, Selbstkonzept" und die Sektion "gesellschaftspolitische Bildung".

Die Aufgaben zur gesellschaftspolitischen Bildung erfordern bloße Begriffskennntnisse, die also nur bedingt dafür geeignet sind, die psychosoziale Entwicklung oder moralische Urteilskraft zu messen.

In der Sektion "Soziale Kompetenz . . ." finden sich bei den Unterkategorien keine weiteren Hinweise auf Fähigkeiten. Wie aber aus der übrigen Berichterstattung über das BIJU-Projekt hervorgeht, sollen diese Fähigkeiten durch die Kategorie "prosoziale Motivation" repräsentiert werden. Es bedarf keiner Spezialisierung um zu erkennen, dass Fähigkeiten nicht dasselbe sind wie Motivationen und umgekehrt. Die Untersuchungsabsichten der Autoren sind also im Messinstrument nur ungenügend, wenn überhaupt repräsentiert.

- Schulen haben die Aufgabe, die *Fähigkeiten* von Heranwachsenden in umfassender Weise zu fördern, also auch im *sozial-moralischen* Bereich und nicht bloß im naturwissenschaftlichen. Hier ist den Autoren der BIJU-Studie entschieden zuzustimmen. An der einseitigen Ausrichtung auf Mathematik und Naturwissenschaften (plus allenfalls Muttersprachenfähigkeit) kranken vor allem die bisherigen internationalen Schulvergleichsstudien (wie TIMSS), an denen teilweise diesselben Autoren beteiligt waren (Baumert, Lehmann et al., 1997).
- Schüler sollen für ihr späteres Leben mit möglichst hohen *Fähigkeiten* - auch im sozial-moralischen Bereich - ausgestattet werden. Ob sie bestimmte moralische *Einstellungen* oder *Werte* haben, kann aber nicht zum Gegenstand von Notengebung oder Evaluation gemacht werden (siehe auch die Kritik von Ratzki, 1998). Es widerspricht unserem *demokratischen Grundverständnis*, dass persönliche *Motive*, Werthaltungen oder Einstellungen zum Gegenstand von Erziehungsprozessen in einer *demokratischen* Gesellschaft gemacht werden. Demnach können solche Motive auch nicht zum Gegenstand der Evaluation von Schulen oder Schulformen gemacht werden. Die Schüler sollen die Aufgaben ihrer Schulfächer möglichst gut beherrschen; ob sie diese auch lieben, muß ihnen überlassen bleiben.
- Der Bewertung, dass Altruismus höherwertig (höher entwickelt) als Egoismus sei, liegt ein sehr subjektives, moralphilosophisch sehr umstrittenes Werturteil zugrunde.⁵ Altruismus kann ebenso ein hohe wie eine niedrige Moralentwicklung anzeigen. Wer immerzu mehr gibt, als von ihm oder ihr gerechterweise erwartet werden kann, verhält sich nicht moralisch, da er/sie das Gleichgewicht von Pflichten und Rechten verletzt, sich selbst also ungerecht behandelt (Krebs, 1982, S. 73 ff.). Andererseits verlangt wohlverstandenes Eigeninte-

⁵ Zur Messproblematik nehme ich weiter unten Stellung.

resse immer auch die Berücksichtigung der Perspektiven anderer und die Sorge für andere (Axelrod, 1981; Simmel, 1989/1892).

Auch "Konformität" kann kaum eindeutig in eine Entwicklungsskala eingeordnet werden, solange man nicht auch erfasst, wie diese Konformität motiviert ist: zum Beispiel durch die Angst vor Vergeltung, die Erwartung von Gratifikationen, das Motiv, soziale Bindungen zu erhalten, oder den Wunsch, einer Norm zu genügen.⁶ Der empirische Nachweis von Alterstrends ist für diese Frage nebensächlich, wenn nicht gar irrelevant.⁷

- Abgesehen davon stellt sich auch die Frage, ob eine solche Studie wichtige Informationen über menschliches Verhalten erbringt. Wie die Altruismusforschung zeigt, haben Motive einen geringen Erklärwert für menschliches Verhalten. Sie korrelieren allenfalls untereinander, was, wie schon lange bekannt ist, auf systematische Tendenzen bei der Selbstwahrnehmung zurückzuführen ist (Mischel, 1968; Cohen, 1969). Zwischen prosozialen *Motiven* und Verhalten bestehen zumeist nur geringe und inkonsistente Korrelationen (Bierhoff, 1990; Krebs, 1970, 1975, 1982; Lind, im Druck).

Die geringe Korrelation zwischen prosozialen Motiven und Verhalten ist zum Teil dadurch bedingt, dass, wie verschiedene Studien zeigen, die *Hilfsbereitschaft* bei Menschen generell sehr hoch ist, es jedoch sehr von individuellen moralischen Kompetenzen (McNamee, 1977) und situativen Bedingungen abhängt, ob diese Bereitschaft sich auch in entsprechendem Verhalten manifestiert. Higgins (1980) fand Hinweise auf die Bedeutung der moralischen Atmosphäre einer Gruppe. "Pluralistische Ignoranz", das heißt die Fehleinschätzung der Hilfsbereitschaft anderer, wirkt offenbar hemmend. Beides wird in der BIJU-Studie nicht untersucht.

Fazit: Die Absicht der Autoren, Fähigkeiten im psychosozialen Bereich zu erfassen, ist in dem Messinstrument der BIJU-Studie nur ungenügend, wenn überhaupt repräsentiert. Prosoziale Motive sind keine Fähigkeiten. Motive eignen sich nicht zur Evaluation von Bildungsprozessen oder Schulformen. Eine solche Evaluation ist nicht vereinbar mit dem Selbstverständnis von Bildung in einer demokratischen Gesellschaft.

⁶ Der Bewertung von *Egoismus* als höherentwickelt und *Altruismus* als niedriger entwickelt liegt wohl eher einer ad hoc-Philosophie zugrunde als eine moralphilosophisch fundierte wissenschaftliche Theorie (wie dies zum Beispiel bei der Theorie der moralisch-kognitiven Entwicklung von Kohlberg, 1985, oder der Altruismus-Theorie von Krebs, 1982, der Fall ist).

⁷ Ich habe am Beispiel der Moralforschung gezeigt, dass das chronologische Alter für die Validierung von Entwicklungsskalen irrelevant ist (Lind, 1993), wie ja auch die Validität physikalischer Messinstrumente (wie das Metermaß) nicht von der tatsächlichen Veränderungen der Gegenstände abhängt, die damit gemessen werden.

Forschungsanlage

Die Autoren bezeichnen die BIJU-Studie wegen ihres Umfangs als “die wichtigste Längsschnittstudie zur Leistungs- und Persönlichkeitsmessung von Jugendlichen im vereinten Deutschland” (MPI Bericht, 1996, S. 5). Im Rahmen der BIJU-Studie wurden zwischen 1991 und 1998 jährlich Schüler und Schülerinnen in neuen und alten Bundesländern befragt. An den drei Einzelstudien (Längsschnittstudie über den gesamten Zeitraum, Querschnittstudie und Nebenkohorte) nahmen anfangs ca. 13.000 Personen teil. Die Längsschnittgruppe verringerte sich von 9.000 in der ersten Befragung auf ca. 5.000 in der 4. Befragungsrunde (MPI Bericht, 1996). Dem Umfang der Erhebungsinstrumente nach zu urteilen hat jede Befragung zwischen zwei und vier Schulstunden erfordert. Über die Kosten für diese Studie liegen keine Informationen vor. Sie waren wohl nicht gering. Allein auf Seiten der teilnehmenden SchülerInnen hat die BIJU-Studie schätzungsweise einen Aufwand von 50.000 bis 100.000 Stunden verursacht.

Kommentar

Die Frage, ob die BIJU-Studie eine der “wichtigsten” Studien ist, wie die Autoren voller Selbstüberzeugung behaupten, kann nicht anhand von Kosten und Umfang entschieden werden. Entscheidend ist dafür vielmehr die Bedeutung der Fragestellung, die Sorgfalt der Untersuchungsplanung und Angemessenheit der Interpretation. Der Kostenaufwand sollte allenfalls Anlass sein, die Maßstäbe für die Güte der Studie besonders streng zu fassen.

Stichproben-Design

Die BIJU-Studie ist als kombinierte Längs- und Querschnittuntersuchung angelegt. Die Auswertungen konzentrierten sich bislang fast ausschließlich auf den Längsschnitt.

Kommentar

- der Hauptvorteil von Längsschnittdesigns ist, dass Unterschiede hinsichtlich der individuellen Ausgangswerte bei der Analyse berücksichtigt und somit Effekt eindeutiger dem Einfluß der Schule zugeschrieben werden können.

- Ihre Hauptnachteile sind, dass durch den Ausfall von Untersuchungsteilnehmern während der Untersuchung (so sank zum Beispiel die Teilnahme zwischen erster und vierter Befragung von 9000 auf 5000 SchülerInnen; siehe MPI Bericht, 1996, S. 11) starke Selektionseffekte wirksam sein können, die einzelne Ergebnisse stark verzerren können⁸; und dass sich die Messungen selbst als Quelle von Veränderungen und somit störend auswirken können. Bei Einstellungs- und Persönlichkeitsfragebogen kann sich vor allem das Phänomen der sozialen Erwünschtheit verzerrend auf die Befunde auswirken (Wiggins, 1973; ich komme auf diese Probleme unten zurück), bei Fähigkeiten die Faktoren Langeweile und Übung.
- Die Autoren der BIJU-Studie haben bei der Analyse der prosozialen Motive bislang weder die Möglichkeit der Längsschnittstudie genutzt, Selektionseffekte zu analysieren und zu kontrollieren (siehe auch die entsprechende Kritik von Anne Ratzki, 1998, S. 21), noch haben sie die bei Längsschnittstudien möglichen Störeffekte ausreichend berücksichtigt. Es ist durchaus möglich, dass die in dem 2. MPI-Bericht (1996) für die Schulen berichteten, schulformtypischen Veränderungen durch starke Selektionseffekte bedingt sind. Die Autoren geben leider nirgends Auskunft darüber, wie die Ausfälle bei der Teilnahme auf die Schulformen verteilt sind, ob neue Teilnehmer "nachgezogen" wurden (was sich bei einer klassenweisen Erhebung anbietet) und wie die Ausfälle mit prosozialen Motiven korrelieren.

Fazit: In der BIJU-Studie werden die Vorteile der Längsschnittforschung, mit denen in der Regel hohe Kosten verbunden sind, bislang nicht genutzt, um alternative Erklärungen für unterschiedliche psychosoziale Veränderungen (wie z.B. Anfangsunterschiede) empirisch zu testen. Auch fehlt bislang eine kritische Analyse der für Längsschnittstudien typischen Verzerrungseffekte (wie Selektion, Ausfälle und soziale Erwünschtheit). So ist nicht klar, ob die Unterschiede der "prosozialen Motivation" zwischen den Schulformen, wie die Autoren meinen, wirklich schulformbedingt ist oder nicht eher durch andere Ursachen, wie z.B. Selektionseffekte oder soziale Erwünschtheit bedingt ist.

⁸ Am Beispiel der Moralentwicklung fand Lind (im Druck), dass durch solche Selektionseffekte bei Längsschnittdesigns der Eindruck einer scheinbaren Höherentwicklung entstehen kann, wo eigentlich eine Regression vorliegt. Mieg (1992, S. 142) fand zudem eine deutliche Korrelation zwischen moralischer Urteilsfähigkeit und Fragebogen-Rücklauf.

Messung “psychosozialer Fähigkeiten” in der BIJU-Studie

Die BIJU-Studie folgt der klassischen Einteilung der Bildungsziele in *kognitive* und *affektive* (Blooms Taxonomie, Bloom et al., 1956; Krathwohl et al., 1964). Als *kognitive* Variablen wurden in der BIJU-Studie Schulleistungstests in fünf Schulfächern (u.a. Mathematik, Physik und Englisch) sowie auf dem Gebiet der allgemeinen Intelligenz erhoben. Als *affektive* Variablen wurden eine Vielzahl von Einstellungen und Selbstbeschreibungen zu den Themen Selbst, Familie, Beruf, Partnerschaft und Gesellschaft erhoben (MPI-Bericht, 1996; Baumert & Köller, 1998). Wie bereits oben erwähnt, enthalten die affektiven Variablen außer Wissensfragen im Bereich der politischen Bildung keine Fähigkeitsaufgaben. Unter der Rubrik “soziale Kompetenz” wurden Selbstbeschreibungen und Einstellungen zu sich selbst und anderen (vor allem Mitschülern) erhoben sowie die “prosoziale Motivation” (BIJU-Erhebungsinstrument; ohne Datum). Die Erhebung prosozialer Motivation wird vor allem damit begründet, dass sie mit der Kompetenz korreliert, die Perspektive anderer einzunehmen und mit anderen mitzuempfinden (Baumert & Köller, 1998, S. 17).

Kommentar

- Nach der Überzeugung von Bloom und seinen Kollegen (Bloom et al., 1956; Krathwohl et al., 1964) kann es so etwas wie “psychosoziale Fähigkeiten” gar nicht geben, da ihrer Taxonomie nach “psychosozial” in die Kategorie “affektive Ziele” gehört und “Fähigkeiten” in die Kategorie “kognitiv”. Vielleicht erklärt dies auch den Umstand, dass es bis heute sehr wenige Ansätze gibt, Fähigkeiten im psychosozialen Bereich zu operationalisieren und zu messen. Eine der wenigen Ausnahmen stellt die moralische Urteilsfähigkeit dar, wie sie von Piaget und Kohlberg definiert wurde.
- Affekt und Kognition sind nur verschiedene Aspekte oder Eigenschaften des Verhaltens. Aber auch wenn sie sich dinglich nicht trennen lassen und manchmal sogar empirisch hoch miteinander korrelieren,⁹ so stellen sie doch zwei

⁹ Jean Piaget und Lawrence Kohlberg haben schon früh im Bereich der Moralentwicklung eine affektiv-kognitive Parallelität angenommen, die auch eine enge empirische Korrelation impliziert. Lange Zeit konnte diese Hypothese aber nicht angemessen untersucht werden, da es mit den vorhandenen Messinstrumenten nicht möglich war, beide Aspekte am selben Verhalten zu messen. Erst mit dem Moralischen Urteil-Test (MUT) von Lind (1978; 1993; 1996) ist es möglich geworden, dies zu tun. Tatsächlich konnte die Parallelitätshypothese damit in vielen Studien eindeutig bestätigt werden. Die Überprüfung ist bislang allerdings fast ausschließlich in Situationen

deutlich unterscheidbare Aspekte des menschlichen Verhaltens dar, die untereinander nicht austauschbar sind. Motive und ähnliche Variablen (wie Einstellungen, Werthaltungen, Präferenzen) sind grundsätzlich in jeder Richtung *simulierbar*, d.h., die Befragten können sie verändern, je nachdem, wie sie instruiert werden oder welche Antworten ihnen sozial opportun erscheinen. Hingegen können sie *Fähigkeiten* nicht beliebig nach oben hin simulieren, falls ihnen dies opportun erscheint.

Fazit: Die Tatsache, dass die Nennung altruistischer Motive mit bestimmten Kompetenzen korreliert, ist kein Beleg dafür, dass die von den Befragten berichteten Motive selbst auch Kompetenzen darstellen.

Die Messung prosozialer Motive

Für die Autoren der BIJU-Studie ist "Schule ein Unternehmen, das unterschiedliche und teilweise konkurrierende Zielsetzungen verfolgt. In gelingenden Bildungsprozessen vollzieht sich [. . .] die gleichmäßige Entwicklung kognitiver, sozialer und emotionaler Fähigkeiten. [. . .] Eingeschlossen ist der Gedanke, dass es Kompensationsmöglichkeiten für geringere Erträge in einem Bereich durch gute Ergebnisse in einem anderen Bereich gebe. Eine zentrale Fragestellung von BIJU ist die Prüfung dieser Annahme. In der Längsschnittstudie wurde deshalb ein breiter Kranz *nicht* [meine Hervorhebung] primär kognitiver Variablen erfasst. [. . .] Für die Erfassung nicht primäre leistungsbezogener Merkmale ist dagegen der Königsweg die Selbstauskunft. [. . .] Solche Selbstauskünfte können trotz aller Antwortverzerrungen relativ zuverlässig mit Fragebögen eingeholt werden" (Baumert & Köller, 1998, S. 15).

erfolgt, die frei von äußerem Druck (Benotung o.ä.) war. Wir können daraus nur schließen, dass Personen mit hohe kognitiven Entwicklungswerten einen moralischen Diskurs auf Stufe 6 und 5 vor allen niedrigeren Stufen bevorzugen. Der Umkehrschluss, dass Leute mit einer Präferenz für höhere Urteilsstufen auch eine hohe Urteilsfähigkeit haben, ist damit nicht zu belegen. Im Gegenteil, Simulationsexperimente widerlegen eine solche Vermutung (Lind, 1996).

- Auch wenn wir nicht akzeptieren, dass es sich bei Motiven, Einstellungen, Werthaltungen etc. um irgendeine Art von Fähigkeit handelt, und wenn wir es für höchst problematisch halten, solche Aspekte menschlichen Verhaltens zum Gegenstand der Evaluation zu machen, so kann ihr Studium doch um ihrer selbst willen interessant und für bestimmte wissenschaftliche Fragestellungen erforderlich sein. Voraussetzung hierfür ist jedoch, dass die Messung dieser Dinge *valide* ist, das heißt, dass mit dem gewählten Messinstrument wirklich das gemessen wird, was die Autoren zu messen beabsichtigen.
- Es reicht für die Frage nach der Validität auf keinen Fall aus zu klären, ob die Messung *zuverlässig* oder *reliabel* ist. So wäre z.B. die Messung der Schuhgröße zwar sehr zuverlässig und reliabel, aber sie würde offenkundig *kein valides* Maß für prosoziale Motive darstellen. Unabdingbar für die Validität einer Messung ist, dass der Messprozess in einer nachvollziehbaren Beziehung zu der Sache steht, die man zu messen beabsichtigt.
- Unter "prosozialer Motivation" versteht man allgemein die Tatsache, dass ein Mensch dazu neigt, anderen zu helfen und im Konfliktfall dazu tendiert, seine eigenen Interessen zugunsten sozialer Interessen zurückzustellen.
- Eine erste Schwierigkeit mit dieser allgemein akzeptierten und scheinbar einfachen Definition besteht darin, dass, wie oben bereits gesehen, das Helfen auch egoistisch motiviert sein kann und daher kein valider Indikator für prosoziale Motivation sein muß. Die Autoren des Fragebogens, Boehnke und Silbereisen, tragen dem Rechnung, indem sie nicht tatsächliches Helfen (oder Helfenwollen) erfassen, sondern die Begründungen der Befragten für solches Verhalten (Boehnke, 1988).
- Allerdings sind auch Begründungen der Befragten für hilfreiches Verhalten *kein* eindeutiger Hinweis auf prosoziale oder altruistische Motive. Diese Begründungen können eine Vielzahl von Dingen anzeigen, die mit dem zu messenden Motiv nichts zu tun haben. Wie schon lange bekannt, ist die Messung von Motiven und ähnlichen Konstrukten (Einstellungen, Werthaltungen etc.) mit vielen Problemen behaftet (Green, 1954; Shaw & Wright, 1967, S. 7 ff.; Emler et al., 1983; Rossi & Anderson, 1982):
 - *Soziale Erwünschtheit* von bestimmten Motiven: Die Befragten geben die Motive an, die ihnen in ihrem eigenen sozialen Kontext sozial oder politisch opportun zu sein scheinen. Für altruistische Motive scheint dies in besonderem Maße zuzutreffen.
 - *Nicht-Existenz von Motiven*: Die Befragten fühlen sich gedrängt, ein Motiv zu wählen, obwohl sie gar keines angeben können oder das wirkliche Motiv ihnen nicht zur Auswahl angeboten wird. Dieses Problem trifft auf den BIJU-Fragebogen zur prosozialen Motivation ganz besonders zu, da hier immer nur drei Motive zur Auswahl stehen.

- *Komplexität von Einstellungen*: Die Motivationsstruktur von Menschen kann weit komplexer sein, als die simple Polarität von Altruismus versus Egoismus insinuiert, worauf bereits Georg Simmel (1892) hinwies. Oft handelt es sich um ein Gemisch mehrerer Motive, die in unterschiedlichen Beziehungen zueinander stehen können. Was gibt ein Befragter an: das Motiv, das ihm zuerst in den Sinn kommt, oder das, das ihm "hinter" diesem Motiv zu stehen scheint, oder das, das seiner Meinung nach überwiegt?
- Das Problem der Validität der Messinterpretation, d.h. die Klärung der oben aufgezeigten alternativen Interpretationsmöglichkeiten erfordert ein sorgfältig geplantes, multivariates Fragebogen-Design (Rossi & Anderson, 1982; Lind, 1996). Davon kann aber bei in der BIJU-Studie verwendeten Fragebogen zur prosozialen Motivation keine Rede sein. Wir haben somit kaum Anhaltspunkte dafür, was dieser Fragebogen eigentlich misst.

Fazit: Selbst wenn man die Messung prosozialer Motivation für sich genommen als ein sinnvolles Unterfangen ansieht (wenn man sie also nicht als Indikator für psychosoziale Fähigkeiten nimmt und nicht zur Evaluation von Schulformen heranzieht), weist die Messmethode der BIJU-Studie schwere Validitätsmängel auf, die eine eindeutige Interpretation der Befunde erschwert, wenn nicht gar unmöglich macht.¹⁰

Nachwort

Für die Messung von psychosozialen Fähigkeiten stehen seit langem Instrumente zur Verfügung, deren Validität eingehend untersucht wurden und die auch schon in umfangreichen Surveys eingesetzt worden sind. Der Referent hat z.B. bereits in den siebziger Jahren ein Instrument zur Messung der moralischen Urteilsfähigkeit ("Urteilstkraft") entwickelt, das allgemein bekannt ist und nachweislich eine Fähigkeit im psychosozialen Bereich misst. Das Instrument, der "Moralische Urteil-Test" (MUT), wurde bereits in mehreren internationalen Forschungsprojekten bei mehreren zehntausend Befragten eingesetzt, auch in Projekten, die die Evaluation von Bildungsprozessen zum Gegenstand haben (siehe Lind, 1978; 1993; 1995).

¹⁰ Die Autoren könnten zwar auf übliche Konventionen hinweisen, aber - in Übereinstimmung mit der Einstufung konventioneller Motive durch die Autoren - müsste man solche Rechtfertigung als entwicklungsbedürftig bezeichnen.

Literatur

- Axelrod, R. (1981). *The evolution of cooperation*. New York: Basic Books.
- Baumert, Jürgen & Lehmann, R., und andere (1997). *Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde*. Opladen: Leske und Budrich.
- Baumert, J. & Köller, O. (1998). Nationale und internationale Schulleistungstudien. Was können sie leisten, wo sind ihre Grenzen? *Pädagogik* 6/1998, 12-18.
- Bierhoff, H. W. (1990). *Psychologie hilfreichen Verhaltens*. Stuttgart: Kohlhammer.
- Bloom, B.S., Engelhart, M.D., Hill, W.H., Furst, E.J. & Krathwohl (1956). *Taxonomy of educational objectives. Handbook I: Cognitive domain*. New York: McKay.
- Boehnke, K. (1988). *Prosoziale Motivation, Selbstkonzept und politische Orientierung: Entwicklungsbedingungen und Veränderungen im Jugendalter*. Frankfurt.
- Cohen, R. (1969). *Systematische Tendenzen bei Persönlichkeitsbeurteilungen. Eine empirische Untersuchung*. Stuttgart: Huber.
- Emler, N., Renwick, S. & Malone, B. (1983). The relationship between moral reasoning and political orientation. *Journal of Personality and Social Psychology*, 45, 1073-80.
- Green, B. (1954). Attitude measurement. In G. Lindzey, Hg., *Handbook of social psychology*, Bd. I. pp. 335-369. Reading, MA: Addison-Wesley.
- Higgins, A. (1980). Research and measurement issues in moral education interventions. In R. Mosher, Hg., *Moral education. A first generation of research and development*, pp. 92-107. New York: Praeger.
- Kohlberg, L. (1995). *Die Psychologie der Moralentwicklung*. Frankfurt: Suhrkamp.
- Krathwohl, D.R., Bloom, B.S. & Masia, B.B. (1964). *Taxonomy of educational objectives. Handbook II: Affective domain*. New York: McKay.
- Krebs, D. (1970). Altruism - An examination of the concept and a review of the literature. *Psychological Bulletin*, 73, 258-302.
- Krebs, D. (1975). Empathy and Altruism. *Journal of Personality and Social Psychology*, 32, 1134-46.
- Krebs, D. (1982). Altruism - a rational approach. N. Eisenberg, Hg., *The development of prosocial behavior*, pp. 53-76. New York Academic Press.
- Lind, G. (1978). Wie misst man moralisches Urteil? In: G. Portele, Hg., *Sozialisation und Moral*, pp. 171-201. Weinheim: Beltz.
- Lind, G. (1995). *The meaning and measurement of moral competence revisited - A dual aspect model*. Invited paper, Annual Meeting of the American

- Educational Research Association, San Francisco, 1995 (<http://www.uni-konstanz.de/ag-moral/mjt-95.htm>).
- Lind, G. (im Druck). Are helpers always moral? Empirical findings from a longitudinal study of medical students in Germany. In U. Gielen, Hg., Human development in cross-cultural perspective.
- McNamee, S. (1977). Moral behavior, moral development and motivation.. *Journal of Moral Education*, 7(1), 27-31.
- Mieg, H. (1992). *Verantwortung. Moralische Motivation und die Bewältigung sozialer Komplexität*. Universität Mannheim: Dissertation.
- Mischel, W. (1968). Personality and assessment. New York: Wiley.
- MPI Bericht (1996). *Bildungsverläufe und psychosoziale Entwicklung im Jugendalter (BIJU)*. 2. Bericht für die Schulen. Berlin: Max-Planck-Institut für Bildungsforschung.
- Ratzki, A. (1998). BIJU und die Gesamtschule. *Pädagogik* 6/98, 20-22.
- Popper, K.R. (1968). Conjectures and Refutations. New York: Harper & Row
- Rossi, P.H. & Anderson, A.B. (1982). The Factorial Survey approach: An introduction. In: P.H. Rossi & S.L. Nock, Hg., *Measuring social judgments: The Factorial Survey approach*, pp. 15-33. Beverly Hills: SAGE Publications.
- Shaw, M. & Wright, J. (1967). *Scales for the measurement of attitudes*. New York: McGraw-Hill.
- Simmel, G. (1989/1892). *Einleitung in die Moralwissenschaft*. (Original 1892) Frankfurt: Suhrkamp.
- Wiggins, J.S. (1973). *Personality and prediction. Principles of personality assessment*. Reading MA: Addison-Wesley.