

# **DOES THE JUST COMMUNITY EXPERIENCE MAKE A DIFFERENCE? MEASURING AND EVALUATING THE EFFECT OF THE DES PROJECT**

by Georg Lind, University of Konstanz, Germany, and Wolfgang Althof, University of Fribourg, Switzerland

## **Introduction**

The question of whether moral education programs make a difference is mostly subject to convictions rather than to empirical research. There are many diverging opinions as to the effects of moral education programs. Many educators want to believe that their interventions make a difference and, therefore, have no doubts about the answer to this question. But only few programs report empirical data by which we could examine their claims.

Among the mostly known programs for fostering moral development that have consistently shown positive effects, is the dilemma discussion program designed by Lawrence Kohlberg, Moshe Blatt and their colleagues (for a review see Leming, 1981; Lockwood, 1978; Schläfli et al., 1985; Lind, 1992). However, the authors of this program themselves argue that its goal is too narrow as it is confined to fostering the individual's moral judgment competence neither than the inter-individual competencies needed to put moral judgment to action in a social context.

Research on the effectiveness of the just community approach is still in its beginning. Power et al. (1989) provide a wealth of information on JC programs in several schools in the United States. But, as the authors indicate, most of this information is anecdotal, and the empirical support is still weak.

When we address the question as to which effects the German just community program has had on students' moral judgment competence we shall have a particular look at the effect-size. Is it larger than the effect-size of regular schooling?

## **Sample and design**

As described in Schirp's article, three schools of different types, located in the German state North-Rhine Westfalia participated in the DES experiment: a lower secondary school ("Hauptschule") with three classes from the sixth grade, a higher secondary school ("Realschule") with two age groups, starting in 1987 and 1988, each comprised of three classes from the fifth grade, and a Grammar School ("Gymnasium") with two classes each from both the sixth and the eighth grade. In total, a number of 395 students took part in the intervention. Most of them have been assessed in the pretest session as well as in the posttest session, data analysis has not been completed with respect to each instrument, however. Depending on local conditions, the testing intervals ranged from 14 to 29 months. Unfortunately, for practical reasons and for reasons of the educational policy in North-Rhine Westfalia it was impossible to match control groups to the interventional groups. Hence, as in most educational experiments, there was no random assignment to experimental and control group. This means that the data presented below provide only a preliminary basis for taking schooling effects into account. In evaluating the results we will have to match our findings to other data representing noninterventional, "natural" developmental processes in the same period of time, &g., data from, the pilot study by Link and Lind (1988) on moral development and moral atmosphere were used for control purposes.

The DES project is being evaluated through a number of both qualitative and quantitative methods like

participatory observations (of lessons and meetings), behavioral analysis of video recordings (especially of die community meetings), in-depth interviews with students and teachers, and several questionnaires.

In the following, we will focus on the results of a Kohlbergian-type written moral judgment interview, of a questionnaire about students' views; of the moral climate of the school, and a responsibility-related test of moral judgment competence. The first questionnaire (i.e., written interview form) explores the moral-cognitive reasoning competence (items referring to two dilemma stories). The second questionnaire is designed to estimate the moral atmosphere in school and classroom and emphasizes the dimensions of participation, helping, rule adherence, and attitudes toward the school, the teachers and schoolmates. The third instrument administered together with the second, is designed to clarify developmental processes in moral judgment competence. Data analysis concerning the second and third questionnaire (conducted by Georg Lind) is complete; results from the first questionnaire (calculated by Wolfgang Althof) are based on a partial sample of 98 students (pro- and posttested), for the time being, and hence are still preliminary - though the findings are so clear that we would be surprised if the overall pattern would change significantly.

- The "Socio-Moral Reflection Measure" (Gibbs, Widamon & Colby, 1982a) that we used in a German language version, adapted to the age group assessed, is a test designed to evaluate the maturity of moral reasoning. It is based on Kohlberg's theory of the development of justice reasoning, and can be understood as an abbreviated version of Kohlberg and Colby's "Moral judgment Interview" (Colby et al., 1987). Given its written form, combined with a stricter standardization in scoring procedures, this test can not only be group-administered but decreases the expenditure of time and money in interviewing and scoring markedly. The SRM was validated in a number of studies (cf. Gibbs et al., 1982a, pp. ME; 1982b). The concurrent validity of the SRM with Kohlberg's MJ is satisfying. Gibbs, Widamon and Colby (1982a, 18) report a correlation of .85 between the two tests, modal stage agreement was 75.4 per cent, with no discrepancies in stage assignments, exceeding one stage. The most important index representing the overall stage ratings of the protocols is the "Sociomoral Reflection Maturity Score" (SRMS) ranging from 100 points (for a clear stage 1) to 400 points (fully elaborated stage 4). For theoretical reasons, there is no assessment of Kohlberg's stage 5; however, this does not affect the utility of the instrument, since stage 5 does not occur before early adulthood.
- The "Moralische Atmosphäre-Fragebogen" (MAP, Lind & Link, 1986) represents an adapted and revised version of Kohlberg, Higgins and Power's Moral Atmosphere Questionnaire (cf. Power et al., 1989), developed by Georg Lind and Lenore Link (1986). It was validated in a study with 216 students (8th through 12th grade). Some parts of the MAF had to be adapted to the verbal skills of 5th graders. The MAF consists of (a) questions pertaining to moral-democratic needs of students (importance of various aspects of the school life), (b) questions concerning different aspects of the school climate (attitudes toward schoolmates, teachers, and the school as a whole; attitudes toward rules at school), as well as a test of the willingness to feel responsible for the following through of something that is considered the right or just thing to do (judgments of responsibility).
- Cases of refusals to fill out the questionnaire have been very rare; obviously, most of the students found it interesting to think about the questions posed. However, some questionnaires were unscorable and had to be eliminated from data analysis because of language problems (students from foreign countries).

## Methodological Issues in Evaluating JC Programs

There are several methodological issues which have not been coped with by recent evaluation research into moral education:

On the one hand, most research is limited to the question whether there is any significant effect. However, the fact that statistically significant differences between a pre-test and a post-test measurement have occurred, tells us little about the educational significance of the particular treatment nor does it support the causal hypothesis that the treatment, rather than other factors, made the difference.

On the other hand, in just community research, there would be several problems in solely measuring moral judgment development by use of Kohlberg's Moral Judgment Interview (MJI). As argued elsewhere (Lind, 1989), the operationalization of moral structure in the MJI is still debated. The scoring of the MJI may be viewed as relying too much on the scorer's ability to "see" consistency in the subjects' reasoning rather than on a transparent scoring procedure which makes direct use of the structural information inherent in the individual's pattern of responses.

- In order to cope with these problems of open interview methodologies, in the DES-project moral judgment was not only assessed with the "Kohlbergian" Socio-Moral Reflection Measure and (for partial samples) with the MR, but also with a new structural test, producing a Moral Judgment Competence Score (M.J.C.S.). While the presentation of this test resembles classical attitude questionnaires, its format and its scoring are designed as a multivariate single subject experiment. The subjects are asked to rate moral arguments according to their acceptability. These arguments represent different stages of moral reasoning as well as the pro and contra opinion on a dilemma. The dilemma is about helping a classmate who is disliked by all students. It is taken from an unpublished moral climate questionnaire by Power, Higgins and Kohlberg. The criterion for getting a high M.J.C.S. score is the degree to which a subjects pattern of responses is morally structured, that is, the degree to which his or her total set of ratings is determined by a moral point of view rather than by other dispositions. The M.J.C.S. theoretically ranges from zero, describing a pattern of ratings which is not at all determined by moral considerations, to one hundred, characterizing a response pattern that is fully determined by a moral point of view.

Note that die M.J.C.S. is not a measure of content - it reflects neither the direction nor the strength of a moral attitude - but a pure measure of the formal property, or structure, of judgment behavior. The M.J.C.S. reflects the degree of response consistency in regard to the moral quality of the stimuli. This is a unique feature of structural scores as compared to content scores like the P-score by James Rest. However, like the P-score the M.J.C.S. does not let us assort subjects into Kohlbergian stages. Note that this scoring of structure is fully transparent and objective, as the scoring does not rely on the work of raters.

The test used here differs in content, but not in design from the more widely published *Moralisches Urteil Test* (MUT) which has been validated in a wide array of research (Lind & Wakenhut, 1985; Lind, 1992). The MUT was not employed in this study because its content is too similar to the written moral judgment interview, also used in this project. The M.J.C.S used here may only be tentatively compared to the MUT scores.

With respect to statistical analyses, we refrain from using statistical significance as a measure of effect-size (see Hildebrand et al., 1977) for two reasons. First, throughout our evaluation of effects, we deal only with populations rather than with randomly drawn samples. Second, as already noted, statistical significance is no adequate measure of effect size since it depends more on the size of the sample than on the amount of effect. Instead we rather use mean score differences which are standardized for the time interval of the intervention.

## Results

### 1. Moral needs of students

In their answers to the MAF, students expressed a strong "moral need" for justice and community in schools, for friendliness and helpfulness among children, for teachers who show fairness and a readiness for substantial dialogues, for good opportunities to learn, for rational rules and adherence to these rules and for participation in decision making processes. Least important they rated the statement that teachers should enforce the schools regulations. Some matters they found not unimportant, but not very important as well, such as being proud of the school, decreasing cheating and competition. This overall pattern of results may seem surprising for many practitioners who underestimate the commitment of students for a school that is a good school both in terms of academic and moral criteria.

After the just community experience - i.e., in post-testing - the list of moral needs remains remarkably stable. Only in a few cases changes in the rating of importance of issues occur. Enforcing school regulation by teachers is perceived even less important. Cheating, school skipping, and fighting became less important problems during the time of the project. In these cases, it is not completely clear if the problems have actually decreased or if the tolerance in face of rule transgressions has increased. No changes can be noted with respect to the need of students for a good academic level, for fair teachers, and for just procedures of decision making in school that include the own participation.

### 2. Moral atmosphere of school

How did the moral atmosphere, the style of social interaction among students and teachers change, as perceived by the students in the MAF? The respective perceptions and opinions of children are not always easy to interpret, as we know from other studies. They are, for example, strongly informed by the given level of cognitive development. The typical pattern is that students have a very strong impression of their school environment at the beginning of their school career. After getting used to it and after their perception becomes more their evaluation gets more and more skeptical or even negative. In a comprehensive study, Fend (1990, p. 101) reports that the contentment with the school declines rapidly between the age of 12 and 15; the number of students who report to feel very content at school decreases from 28 per cent to 7 per cent.

The DES students show only a minor decrease in positive evaluations of their schools. With respect to two issues the positive judgment remained particularly stable: their teachers' behavior and the opportunities to participate in decision making are still valued very high. In one of the schools we had a particular opportunity to check if the experiment had a positive influence on the moral atmosphere of the school. In the higher secondary school at Langenfeld two age groups participated in the intervention. The just community was started some months after the first group had entered the school; the second group joined the school after the just community had already worked for a while. The 5th graders of 1988 (i.e., the second group) rated the same school markedly more positive than the 5th graders of 1987 with respect to almost each issue: (a) They have the impression that teachers treat students more respectfully, are more open, more encouraging, and more discursive. Even the lessons are perceived as better than the year before. (b) After the first year of the just community project the school is more clearly perceived as a community with people helping one another and following common rules. (c) Students see better opportunities to participate in decision making.

### 3. Moral reasoning competence

**Table 1: Changes in moral judgment level (SRMS) in 3 subsamples (n = 98)**

	Pre (means)	Post (means)
1987 age group of "Realschule"	203	243
1988 age group of "Realschule"	222	241
First age group of "Gymnasium"	235	278

Factor	df	F	p
Age group	2	14.35	.000
Time	1	105.79	.000
Age group by time	2	6.11	.003

In the data analyzed thus far, a strong increase in moral reasoning competence (as measured by the Sociomoral Reflection Measure) shows up. The results that can be reported here cover the two age groups in the "Realschule" (6 classes, two thirds of the pretest and posttest protocols analyzed) and the younger of the two age groups in the High School (half of the protocols analyzed). The following picture emerges (see table 1):

The 1987 age group in the higher secondary school shows an increase of 40 SRMS points (203 to 243) during a 23 month period of time. The 1988 age group in the same school scored from a higher level of moral judgment maturity and ends up at the same level as the other group; the gain is 19 points in 15 months (222 to 241). The two Grammar School classes show the most marked increase in moral judgment (43 SRMS points), although they were on a significantly higher level from the very beginning (change from 235 to 278 points). All time effects are highly significant ( $p = .000$ ).

Let us have a look at a comparison of genders. In their studies, Gibbs, Widamon and Colby reported no significant sex differences. Our data confirm this finding in a most clearly way. The curves for girls and boys are absolutely identical (cf, Table 2). The overall picture does not change at all when we have a more differentiated look at the individual age groups or school classes: in no case, pretest or posttest, the difference between boys and girls is higher than 9 SRMS points, in most cases it does not exceed 2, 3 or 4 points - differences that are, of course, irrelevant, both statistically and theoretically.

The general increase just reported represents a transition in moral judgment competence that is of extreme importance, both psychologically and practically. In the "Realschule" the moral thinking of most students was - more or less - pure Stage 2. Stage 2 is the stage of the concrete and individualistic understanding of justice that

**Table 2: Changes in moral judgment level (SRMS) by gender**

	Pre (means)	Post (means)
Female (n = 41)	217	249
Male (n = 57)	218	249

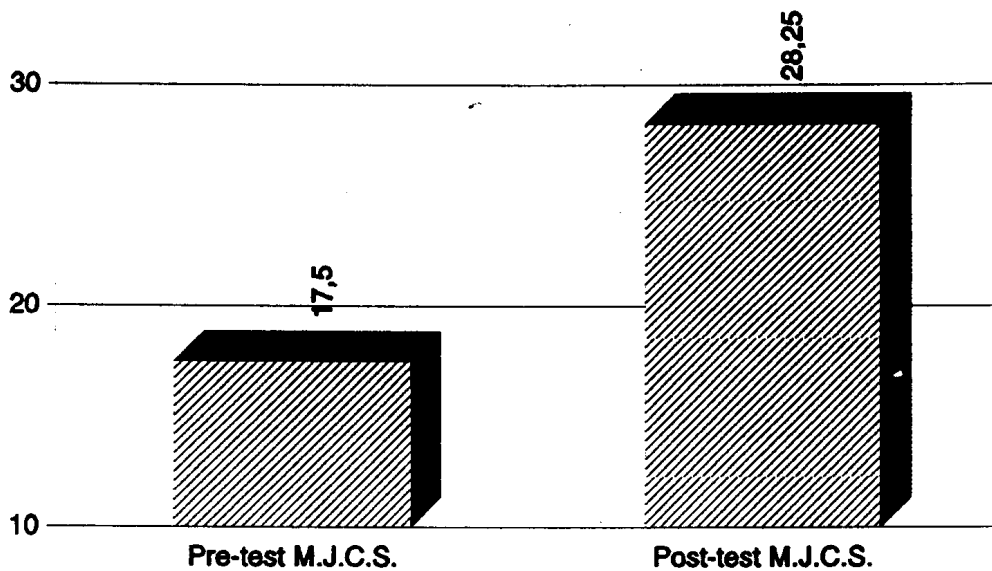
  

Factor	df	F	p
Sex	1	.02	.891
Time	1	87.62	.000
Sex by time	1	.02	.894

requires equal returns for each service and each harm. In the second assessment, after one to two years of just community experiences, the students are clearly on their way to Stage 3, the stage of the morality of the group requiring consideration of the other person's needs, the stage of a conception of justice that operates with

Fig. 1

### Mean M.J.C.S. Development (Test: Responsibility Judgment)



DES-project; N= 395  
• 1992 by G. Lind

notions like reliability, trust, loyalty, and commitment to a common endeavor. That is, they are on their way (and many of them already made it) to a thinking that for the first time makes it possible in a comprehensive sense to view the school as a community and not only as a crowd of individuals looking out only for him- or herself. The same pattern holds true and is even more pronounced in the Grammar School. The students (aged 12 years at the beginning, and 13 ½ years by time of the second assessment) have become Stage 3 on the average, with a considerable number of boys and girls beginning to head for Stage 4.

We have not yet conducted a systematic calculation of the specific effects of the DES program, as compared to "natural" developmental changes in moral reasoning ability in a comparable period of time. This will be done when the data will be complete. For the time being, we can refer to cross-sectional comparisons of age groups, as reported by Gibbs, in order to gain an impression of natural age trends: In Gibbs initial study in the United States, the mean reasoning of the 5th graders showed 202 SRMS points, the 7th graders' average was at 235 SRMS points (see Gibbs et al. 1982a, p. 19; 1982b, p. 907). The level of moral reasoning at die "Realschule" is fairly similar, the level of the Grammar School sample is higher, the increase is both cases is higher than in Gibbs' cross-sectional sample, but not markedly so.

**Table 3: Mean M.J.C.S. Development by School**

School	Total N	Grade	DES project		Mean change	Time (months)	Effect Size in 1 Year
			Pre-test M.J.C. score	Post-test M.J.C. score			
"Hauptschule"	100	6th/8th	22,00	41,00	19,00	29	7,86
"Realschule" (first year)	89	5th/7th	14,88	32,50	17,62	24	8,81
"Realschule" (second year)	85	5th/6th	13,50	17,25	3,75	14	3,21
"Gymnasium"	121	6th/7th 8th/9th	20,67	28,00	7,33	20	4,40
<b>All Schools</b>	<b>395</b>		<b>17,50</b>	<b>28,25</b>	<b>10,75</b>	<b>21,75</b>	<b>5,93</b>
<b>Estimation of the Effect of Schooling</b>							
<b>Pilot study</b>	<b>219</b>	<b>8th/11th</b>	<b>30,75</b>	<b>41,50</b>	<b>10,75</b>	<b>36</b>	<b>3,58</b>

#### 4. Dependent Measure: Moral Judgment Competence Score M.J.C.S. (Responsibility Judgment)

What do the M.J.C.S. data tell us concerning our first main questions, i.e., whether any moral development took place during the DES project?

The data show that, again, there are considerable over-all differences between pre-test and post-test. The median M.J.C. scores are 17.5, mid 28.25 respectively, yielding a development of 10.75 points for the total intervention (see Fig. 1). A score of 10.75 means that the intervention effect is as large as ten percent of the absolute scale range (0 to 100 M.J.C.S. points). This finding does not yet take into account the fact that at average the intervals between pre-tests and post-tests were 21.75 months. The standardized one-year median M.J.C.S. development during the DES project was 5.93. This is still a considerable effect size.

Our second question regarded the degree to which this effect can be attributed to the educational intervention rather than to other factors that may have affected moral development during this time. Most authors mention age as an important factor. Yet, as Jim Rest (1986) has aptly argued, chronological age does not play an important role in moral development by itself. The main factor in advancing moral judgment competence is the amount of general schooling that a person gets. When schooling ceases early in life moral judgment competence regresses (see also Lind, 1992).

Our analysis of the effect of the DES project shows a clear intervention effect irrespective of any schooling effect. On the basis of a pilot study the average schooling effect on M.J.C.S. development is 3.58 per year (see Table 3). Taking the schooling effect into account, gives a pure intervention effect of 2.35 M.J.C.S. points (= 5.93 minus 3.59). Hence, the just community program almost doubled the effect of normal schooling on moral development.

## Discussion

The area of moral education has long been dominated by beliefs and contentions, and only recently the path-breaking cognitive-developmental approach to child and adolescent psychology by Lawrence Kohlberg and his associates has made it possible to evaluate intervention effects in the moral domain on an empirical basis.

On the basis of our findings we conclude that the Just Community approach to moral education, as outlined by Kohlberg, Clark Power, Ann Higgins, and as adapted and extended by Fritz Oser and others, has a considerable effect for the development of moral judgment competence. It seems that schools can considerably enlarge their educational effectiveness when the just community approach to moral education is employed. The success is not dependant on the school type (but it certainly depends on a number of preconditions within the school, as Fritz Oser showed in his introduction to this issue). A combination of moral education in the classroom and democratic participation in school affairs can effectively foster the moral and democratic reasoning competence of students and the moral atmosphere of schools.

The effects represented in our preliminary results go beyond the normal effects of schooling, and - as such - they are not tied to particular types or levels of schools. The students at lower levels in the system of educational stratification show as clear gains in their reasoning competence as students at higher levels. Classroom dilemma discussions and the participation in community meetings and activities helped the students to perceive and judge their school environment clearer and in a more differentiated way. Most of all, it helped them to orient themselves stronger to moral norms and principles when dealing with problematic issues. Students became increasingly capable to examine opinions under rational and moral points of views.

However, our findings are still preliminary in many respects. In order to provide a basis for reliable information we need more data from carefully designed intervention programs. Among other things, we need control groups which match the experimental groups in regard to years of schooling and school type; the intervention intervals should be varied more systematically; post-post-tests should be administered to measure the endurance of effects; objective measures for other outcomes of moral education should also be developed.

Many interesting questions are still open for further inquiry. For example, many teachers wonder whether or not their efforts can have any effect at all. This may be because many teachers are unfamiliar with slow, long-term and seemingly insignificant developmental processes. This may lead to impatience and a premature focus on more easily measurable instructional and behavioral effects. Evidently, in order to keep up motivation over such long intervention intervals, this suggests at least two consequences: The need for more detailed process evaluation and continuous feedback (eg., through intermediate testing of moral judgment competence) and the need for a thorough preparation, in-service training and supervision aimed at empowering teachers, at helping them to improve their understanding of students' developmental achievements and potentials and their instructional, educational, and human capabilities.

Note by the editor of this re-print (2002):

The original text has been scanned and OCRed; so the pages are not the same as in the original.

More data on this intervention project can be found in these publications:

Lind, G. (2002). *Ist Moral lehrbar? Ergebnisse der modernen moralpsychologischen Forschung*. Berlin: Logos.

Lind, G. (2002). Web-site "Psychology of Morality and Democracy & Education",

<http://www.uni-konstanz.de/ag-moral/>



## References

- Colby, A. / Kohlberg, L., et ai. (1987). The Measurement of Moral Judgment. Volume 1: Theoretical Foundation and Research Validation. Volume 11. Standard Issue Scoring manual. Cambridge (England) etc.: Cambridge University Press.
- Dobbelstein-Osthoff, P. /Lind, G. / Oser F./Reinhardt, S./ Schirp, H. (1991). Demokratie und Erziehung in der Schule. Förderung moralisch-demokratischer Urteilsfähigkeit. Bericht des Landesinstituts für Schule und Weiterbildung. Soest.
- Fend, H. (1990). Vom Kind zum Jugendlichen. Bern: Huber.
- Gibbs, J. C. / Widamon, K. F. / Colby, A. (1982a). Social Intelligence. Measuring the Development of Socio-moral Intelligence. Englewood Cliff, NJ.: Prentice Hall
- Gibbs, J. C. / Widamon, K.F. / Colby, A. (1982b). Construction and validation of a simplified, group-administerable equivalent to the Moral Judgment Interview. *Child Development*, 53, 895-910.
- Hildebrand, D. K. / Laing, J. D. / Rosenthal, H. (1977). Prediction Analysis of Cross New York: John Wiley & Sons.
- Kohlberg, L (1985). Foreword. In G. Lind, H. A. Hartmann & R. Wakenhut, eds., Moral development and the social environment. Studies in the philosophy and psychology of moral judgment and education, pp. xv-xvii Chicago. Precedent Publishing Inc.
- Leming, J. S. (1981). Curricular effectiveness in moral values education: A review of research. *Journal of Moral Education* 10, 147-164.
- Lind, G. / Raschert, J., eds., (1997). Moralische Urteilsfähigkeit. Eine Auseinandersetzung mit Lawrence Kohlberg über Moral, Erziehung und Demokratie. Weinheim: Beltz.
- Lind, G. / Wakenhut, R. (1985). Testing for moral judgment competence. In G. Lind, H. A. Hartmann & R. Wakenhut, eds., Moral development and the social environment. Studies in the philosophy and psychology of moral judgment and education, pp. 79-105. Chicago: Precedent Publishing Inc.
- Lind, G. (1989). Measuring moral judgment. A review of the Measurement of Moral Judgment by Anne Colby and Lawrence Kohlberg. *Human Development* 32, 389-397.
- Lind, G. (1990) Wirksamkeit von Ansätzen zur Förderung der moralisch-demokratischen Urteilskompetenz nach Lawrence Kohlberg. In: U. Steffens, T. Bargel. Arbeitskreis "Qualität von Schule". Heft 7: Erziehung und Demokratie in der Schule - Konzepte und Erfahrungen zum Ansatz- von Kohlberg und Oser. Hessisches Institut für Bildungsplanung und Schulentwicklung. Wiesbaden 1991.
- Lind, G. (1992). Moralische Entwicklung und Bildung. Heidelberg: Asanger (in press).
- Lind, G. / Link, L (1986). Moralische Atmosphäre Fragebogen (MAF). Theoretische Grundlagen und Instrument. University of Konstanz.
- Lind, G. / Link, L. (1988), Entwicklung der sozialen Verantwortung in der Schule. Studien zur moralisch-demokratischen Urteilsfähigkeit. Unpublished paper, University of Konstanz.
- Lockwood, A. L. (1978). The effects of value clarification and moral development curricula on school-age subjects: a critical review of recent research. *Review of Educational Research* 48, 325-364.
- Oser, F. / Althof, W. (1992). Moralische Selbstbestimmung. Modelle der Erziehung und Entwicklung im Wertebereich. Stuttgart: Klett-Cotta.
- Power, C. / Higgins, C. / Kohlberg, L. (1989). Lawrence Kohlberg's Approach to Moral Education. New York: Columbia University Press.
- Rest, L. (1986). Moral Development. Advances in Research and Theory. New York: Praeger.
- Schläfli, A. / Rest, J. R. / Thoma, S. J. (1985). Does moral education improve moral judgment? A meta-analysis of intervention studies using the Defining Issues Test. *Review of Educational Research* 55, 319-352.